```
GenCore version 5.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 19:47:58; Search time 15.5 Seconds

(without alignments)

94.913 Million cell updates/sec

Title: US-09-725-010-2

Perfect score: 279
Sequence: 1 CTNCGTKTTPLWRRNPQGQP......LHGVVRPLSLKTDVIKKRQR 50

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574
```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
4: /cgn2\_6/ptodata/1/iaa/Ba\_COMB.pep:\*
5: /cgn2\_6/ptodata/1/iaa/PcTUS\_COMB.pep:\*
6: /cgn2\_6/ptodata/1/iaa/PcTUS\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

|           |                | Annli           | Annii           | Appli            | Appli           | Appl             | Appl              | Appl             | Appl             | Appl              | Appli           | Appl i          | Appl             | Appli            | Appli            | Appli           | Appli           | Appli           | Appl             | Appl             | Appl             | Appli           | Appl     | Appl     | Appl          | App1     | , Appl   | Appl              |
|-----------|----------------|-----------------|-----------------|------------------|-----------------|------------------|-------------------|------------------|------------------|-------------------|-----------------|-----------------|------------------|------------------|------------------|-----------------|-----------------|-----------------|------------------|------------------|------------------|-----------------|----------|----------|---------------|----------|----------|-------------------|
|           | ü              | - 6             | ì               | ,                | ١,              | 11               | 1                 | H                | Ξ                | Ξ                 | _               | ì               | 15               | 4                | 7                | m               | 7               | ìκ              | 27               | 27               | 27               | ω               | 1        | 17       | 37            | 37,      | 37,      | 36,               |
|           | Description    | Segment         | Segmente        | Segmence         | Sequence        | Sequence         | Sequence          | Sequence         | Sequence         | Sequence          | Sequence        | Sequence        | Sequence         | Sequence         | Sequence         | Sequence        | Sequence        | Sequence        | Sequence         | Sequence         | Sequence         | Sequence        | Sequence | Sequence | Seguence      | Sequence | Sequence | Sequence          |
| SUMMARIES | ΩI             | US-09-254-325-2 | US-09-079-415-6 | US-08-750-458A-2 | US-08-927-394-2 | US-08-204-740-11 | US-09-081-167A-11 | US-09-081-395-11 | US-09-416-833-11 | PCT-US95-02521-11 | US-09-008-465-1 | US-09-528-959-1 | US-08-117-083-61 | US-09-362-123A-4 | US-08-269-441A-2 | US-09-008-465-3 | US-09-276-851-2 | US-09-528-959-3 | US-08-911-853-27 | US-09-479-409-27 | US-09-479-453-27 | US-08-852-153-8 | -08-39   | -09-56   | -08-162-081B- | -780-    | -09-082- | US-08-162-081B-36 |
|           | DB             | : 6             | m               | m                | ~               | _                | ٣                 | ٣                | 4                | Ŋ                 | 4               | 4               | Н                | 4                | -                | 4               | 4               | 4               | ٣                | 4                | 4                | 7               | m        | 4        | ~             | 7        | 4        | (7)               |
|           | Length         | 853             | 866             | 866              | 449             | 108              | 108               | 108              | 108              | 108               | 341             | 341             | 416              | 543              | 314              | 314             | 314             | 314             | 381              | 381              | 381              | 664             | 1068     | 1068     | 1069          | 1069     | 1069     | 1080              |
| æ         | Query<br>Match | 92.5            | 92.5            | 92.5             |                 | 19.4             | 19.4              | 19.4             | 19.4             | 19.4              | 18.5            | 18.5            | 18.3             | 18.3             | 18.1             | 18.1            | 18.1            | 18.1            | 18.1             | 18.1             | 18.1             |                 | ٠        |          | •             | 17.9     | •        | 17.9              |
|           | Score          | 258             | 258             | 258              | 188             | 54               | 54                | 54               | 54               | 54                | 51.5            |                 | 51               | 51               | 50.5             | 50.5            | 50.5            | 50.5            | 50.5             | 50.5             | 50.5             | 20              | 20       | 20       | 20            | 20       | 50       | . 50              |
|           | Result<br>No.  | 1               | 7               | ٣                | 4               | 'n               | φ                 | 7                | 8                | 6                 | 10              | 11              | 12               | 13               | 14               | 15              | 16              | 17              | 18               | 19               | 20               | . 21            | 22       | 23       | 24            | 25       | 26       | 17                |

| Sequence 36, Appl<br>Sequence 3, Appl<br>Sequence 1, Appli<br>Sequence 1, Appli<br>Sequence 14, Appli<br>Sequence 12, Appl<br>Sequence 52, Appl<br>Sequence 52, Appl<br>Sequence 52, Appl<br>Sequence 52, Appl<br>Sequence 6, Appli<br>Sequence 6, Appli<br>Sequence 6, Appli<br>Sequence 2, Appli<br>Sequence 2, Appli<br>Sequence 2, Appli<br>Sequence 2, Appli<br>Sequence 2, Appli<br>Sequence 2, Appli  |
|--|
| US-08-780-872-36<br>US-09-085-957-36<br>US-09-579-181-1<br>US-09-134-0010-4442<br>US-09-135-166-1<br>PCT-US95-06266-141<br>US-08-974-022-52<br>US-08-974-186-52<br>US-08-974-186-52<br>US-08-975-447A-52<br>US-08-975-4486-52<br>US-08-975-4486-52<br>US-08-195-448A-6<br>US-08-150-86AA-6<br>US-08-150-86AA-6<br>US-08-120-86AA-6<br>US-08-012-269A-2<br>US-08-012-269A-2<br>US-08-012-269A-2<br>US-08-012-269A-2<br>US-08-012-269A-2<br>US-08-012-269A-2<br>US-08-012-269A-2 |
| N444440004444444444  |
| 1080<br>2972<br>3118<br>249<br>249<br>3118<br>249<br>191<br>191<br>191<br>191<br>191<br>191<br>191<br>256<br>256<br>256<br>256<br>256<br>256   |
| 000000000000000000000000000000000000000  |
| 117777777777777777777777777777777777777  |
| 44<br>0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  |
| 00000000000000000000000000000000000000   |
| ,  |

# ALIGNMENTS

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .<u>;</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ouery Match 92.5%; Score 258; DB 3; Length 853; Best Local Similarity 92.0%; Pred. No. 8.6e-29; Matches 46; Conservative 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              652 CTNCFTQTTPLWRRNPEGQPLCNACGLFLKLHGUVRPLSLKTDVIKKRNR 701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "TRUNCATED AREA WHICH IS
STILL ACTIVE BUT NOT REPRESSED BY
L-GLUTAM..."
                                                                                                                                                                                                                                                   SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30 (EPO)
                                                                                                                           ENHANCED EXPRESSION OF PROTEOLYTIC ENZYMES IN KOJI MOLDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LOCATION: 652-676
OTHER INFORMATION: /note= "DNA BINDING SITE"
                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/254,325
                                     Sequence 2, Application US/09254325
Patent No. 6090607
GENERAL INFORMATION:
                                                                                                                                                                                                                               COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                         FILING DATE:
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                       TITLE OF INVENTION: ENHANCI
TITLE OF INVENTION: PROTEOI
NUMBER OF SEQUENCES: 17
COMPUTER REALABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 853 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY: Binding-site
LOCATION: 652-676
                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE CHARACTERISTICS:
LENGTH: 853 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCATION: 1.731

COTHER INFORMATION:
COTHER INFORMATION:
COTHER INFORMATION:
US-09-254-325-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY: Region
                                                                                                                                                                                                                                                                           SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                   US-09-254-325-2
                                                                                                       APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURE
RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
```

Db 652 CTNCFTQTTPLWRRNPEGOPLCNACGLFLKLHGVVRPLSLK RESULT 2 US-09-079-415-6 Sequence 6, Application US/09079415 Patent No. 6013452 GENERAL INFORMATION: 4129.204-US

```
6 08:17:52 2003
```

```
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY:
                                                                                                                                                                                              US-08-750-458A-2
                                                       TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-927-394-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-927-394-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 4
                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
APPLICANT: Christensen, Tove
APPLICANT: Hynes, Michael J.
TITLE OF INVENTION: A Fungus Wherein The AreA Gene Has Been Modified And An AreA
NUMBER OF SEQUENCES: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60251850 No. 6025185disk of No. 6025185th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60134520 No. 6013452disk of No. 6013452th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.5%; Score 258; DB 3; Length 866; 92.0%; Pred. No. 8.7e-29; tive 2; Mismatches 2; Indels
            APPLICANT: Lehmbeck, Jan
TITLE OF INVENTION: A Fungus Wherein The area, pepC and/or
TITLE OF INVENTION: pepE Genes Have Been Inactivated
NUMBER OF SEQUENCES: 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  664 CTNCFTQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKIDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: DOS SOFTWARE: FastSEO for Windows Version 2.0 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4657.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICATION NUMBER: US/08/750,458A
                                                                                                                                                                                                                                                                                                                                                      UMBER: US/09/079,415
14-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/08750458A Patent No. 6025185
                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4657
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3-December-1996
                                                                                                                                     STREET: 405 Lexington Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36,993
  APPLICANT: Christensen, Tove
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATTORNEY/AGENT INFORMATION:
NAME: ROZEK, Carol E.
REGISTRATION NUMBER: 36,9
                                                                                                                                                                                                                                                                            IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 866 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                     COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER:
FILING DATE: 14-MAY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : New York
RY: U.S.A.
10174-6401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 3-CLASSIFICATION:
                                                                                                                                                           New York
                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New York
                                                                                                                                                                                                USA
                                                                                                                                                                                                                    10174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                              COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-750-458A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY:
                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-079-415-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                         STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
```

```
TITLE OF INVENTION: GATA-6 TRANSCRIPTION FACTOR: COMPOSITIONS AND METHODS NUMBER OF SEQUENCES: 10 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
()
                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                       92.5%; Score 258; DB 3; Length 866; 92.0%; Pred. No. 8.7e-29; tive 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                       664 CINCFIQITPLWRRNPEGOPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 2; Length 449;
                                                                                                                                                                                                                                                                                                                              1 CINCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOFTWARE: Patentin Release #1.0, Version #1.25 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67.4%; Score 188; DB 2; 62.0%; Pred. No. 5e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wolf, Greenfield, & Sacks, P.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S1237/7005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/025,574
FTLING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: US/08/927,394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application US/08927394 Patent No. 5990092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: S12
LECOMMINITATION - S12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 600 Atlantic Avenue
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Plumer, Elizabeth R. REGISTRATION NUMBER: 36,63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 866 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    449 amino acids
                                                                                                                                                                                                                                              Query Match 92.55
Best Local Similarity 92.03
Matches 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               617-720-2441
                                           212 867 0298
                                                                                                                                                  TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE TYPE: protein
```

```
09c414 aspergillus
09y768 aspergillus
09haw7 cladosporiu
060043 metarhizium
P87016 zygosacchar
09p552 schizophyll
059442 penicillum
09hev2 emericella
09hev4 emericella
09hev4 emericella
09hev4 emericella
09hev4 emericella
09hv4 emericella
09hv5 anopheles
095vy5 anopheles
09ty5 aropheles
09ty18 (cratifis c
08wqq4 calliphora
095vy5 anopheles
09ty18 (crosophila
                                                                                                                                       January 3, 2003, 19:42:57; Search time 48 Seconds (without alignments) 214.632 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                          US-09-725-010-2
279
I CINCGIKTIPLWRRNPQGQP......LHGVVRPLSLKTDVIKKRQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        671580
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                     671580 segs, 206047115 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         099C414
099C418
099YEB
099YEB
060043
099052
09982
099HEVZ
099H
                                                                                                                                                                                                                                                                                                                                           BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_invertebrate:*
sp_mammal:*
sp_mhc:*
sp_organelle:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_virus:*
sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_archea:*
sp_bacteria:*
sp_fungi:*
sp_human:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_rvirus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SPTREMBL_21:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                                                                                                                                                                                              Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  258
258
2258
2242
2203
2203
2203
2201
199
198
198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                                                             Sequence:
                                                                                                                                              Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Title:
```

| 17   191   68.5   95.0   5   99VD1   099VE01   099VE01   077157   51Concol   18   67.4   191   6   6   955AA5   057157   057157   51Concol   18   67.0   338   13   059VE01     | 11111111111111111111111111111111111111  |
|---|---|
| 00000000000000000000000000000000000000  | 11111111111111111111111111111111111111  |
| 00000000000000000000000000000000000000  | 11111111111111111111111111111111111111  |
| 00000000000000000000000000000000000000  | 11111111111111111111111111111111111111  |
| 11111111111111111111111111111111111111  | 11111111111111111111111111111111111111  |
| 11111111111111111111111111111111111111  | 11111111111111111111111111111111111111  |
| D O O WELL OF THE PROPERTY OF | S S P P P P P P P P P P P P P P P P P P |
|   |   |

```
g
                                                                   ŏ
                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN=ATCC 56775;
MEDLINE=20225677; PubMed=10760588;
Chang P.-K., Yu J., Bhatnagar D., Cleveland T.E.;
"Characterization of Aspergillus parasiticus major nitrogen regulatory
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Expression of the avirulence gene Avr9 of the fungal tomato pathogen Cladosporium fulvum is regulated by the global nitrogen response
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                             Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=5067;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Fungi; Ascomycota; Pezizomycotina;
Dothideomycetes et Chaetothyriomycetes incertae sedis;
Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Perez-Garcia A., Snoeijers S.S., Joosten M.H.A.J., Goosen T.,
De Wit P.J.G.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 866;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.5%; Score 25%; DB 3; Length 86 92.0%; Pred. No. 2.1e-26; Live 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HIGH TO THROUGH TO THE TOTAL TO THE TOTAL THROUGH THE TOTAL THROUGH TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HOSPITE TATA AND THE TOTAL EXTERS TO THE TOTAL AND THE TOT
                                                      01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
866 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              918 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gene, area.";
Balcohim. Blophys. Acta 1491:263-266(2000).
EMBL; AF148539; AAD37409.1; -.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mol. Plant Microbe Interact. 0:0-0(2001).
EMBL; AF312694; AAG48616.1; -.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nitrogen response factor NRF1. Cladosporium fulvum (Fulvia fulva).
                                                                                                                                                   Major nitrogen regulatory protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                Aspergillus parasiticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-5499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            factor NRF1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
09Y7E8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9HEW7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9HEW7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09HEW7
   à
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae; Metarhizium. NCBI_TaxID=5530;
                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
Ushio K., Otsuka H., Yoshikawa S., Taguchi G., Shimosaka M.,
Mitsui N., Okazaki M.;
"Cloning of the SATI gene concerned with salt tolerance of the yeast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Screen S.E., Bailey A.M., Charnley K., Cooper R., Clarkson J.M.; "Isolation of a nitrogen response regulator gene (nrrl) from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zygosaccharomyces rouxii (Candida mogii).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  663 CINCETQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 712
                                                                                                                           673 CINCFIQITPEMRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92.5%; Score 258; DB 3; Length 944; 92.0%; Pred. No. 2.3e-26; Live 2; Mismatches 2; Indels
     Length 918;
                                                                                                  20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Metarhizium and misopliae.",
Ligene 221:17-24(1998).
EMBL; AJ006468; CAA07052.1; -.
EMBL; AJ006468; CAA07052.1; -.
RISSP, PI7429; AGAT.
RICEPPO: IPR000679; Znf_GATA.
RICHER, PR00619; GATA.1: -.
REMINIS: PR00619; GATA.1: -.
REMINIS: PR00619; GATA.2N.FINGER.
REMINIS: PR00411; Znf_GATA.1: -.
RPROSITE; PS00144; GATA_ZN.FINGER_1; 1.
RPROSITE; PS50114; GATA_ZN.FINGER_2; 1.
SEQUENCE 944 AA; 99687 MW; AlA723E658C23EA0 CRC64;
                                                                                                  1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKIDVIKKRQR
                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 07, Last sequence update) Nitrogen response regulator.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
Score 258; DB 3; L
Pred. No. 2.2e-26;
2; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    327 AA.
                                                  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            J. Ferment. Bioeng. 82:16-21(1996).
EMBL; D83211; BAA18948.1; -.
EMSSP; P17429; 4GAT.
InterPro; IPRO00679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-ME1;
MEDLINE-99070047; PubMed-9852945;
     92.5%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUL-1997 (TrEMBLrel. 04, 01-JUL-1997 (TrEMBLrel. 04, 01-JUN-2001 (TrEMBLrel. 17, DNA-binding protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match
Best Local Similarity 92.09
Matches 46; Conservative
                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zygosaccharomyces rouxii.
                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Metarhizium anisopliae.
                           Best Local Similarity
Matches 46; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=4956;
       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P87016;
                                                                                                                                                                                                                                                                        060043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P87016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
                                                                                                                                                                                                                        RESULT 4
                                                                                                                                                                                                                                                 060043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P87016
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                   Q9HEV2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09HEV3;
                                                                                                                                                                                                                                                                                                                                                                             Q9HEV2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9HEV3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AREB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AREB.
                                                                                                                                                                                                                                                                                                                                RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
                                                                                                                                                                                                                                                                                                                                                      Q9HEV2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        09HEV3
         SORBER
                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                   ACCOCCOS DE LA COCCO COC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schizophyllum commune (Bracket fungus).
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Schizophyllaceae; Schizophyllum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota, Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes,
Eurotiales, Trichocomaceae, mitosporic Trichocomaceae, Penicillium,
NCBI_TaxID=5076;
                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        957 CINCQITNIPLMRRDPEGQPLCNACGLFYKLHGVVRPLSLKIDVIKKRNR 1006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 1075;
                                                                                                                                                                                                                                          11:1 1:11|1|1|1|1|1|246 CTHCHTRTTPLWRRDPMGNPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 295
                                                                                                                                  Length 327;
                                                                                                                                                                                                                   CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                Score 248; DB 3; Length 32
Pred. No. 1.8e-25;
3; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00320; GATA; I...PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZnF_GATA; I.
PROSITE; PS00344; GATA_ZN_FINGER_1; I.
PROSITE; PS0114; GATA_ZN_FINGER_2; I.
SEQUENCE 1075 AA; 113322 MW; D29343211C02FFEA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yamagishi K.;
"putalive transcription factor ScGATA-6.";
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AB043467; BAA96108.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Haas H., Angermayr K., Zadra I.;
Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
EMBL, 196385, AAC09045.1; --
HSSP: P17429; 4GAT.
InterPro; IRR006679; Znf_GATA.
Pfom; PF00320; GATA; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR
                      SMART; SM00401; ZNF_GATA; 1.
PROSTTE; PS50114; GATA_ZN_FINGER_2; 1.
DNA-binding 327 Aa; 37304 MW; 9D7AF1002807283F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-UN-2002 (TrEMBLrel. 21, Last annotation update)
Transcription factor ScGATA-6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1998 (TrEMBLrel. 07, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86.7%; Score 242; DB 3;
86.0%; Pred. No. 3.8e-24;
tive 2; Mismatches 5.
                                                                                                                                                                                                                                                                                                                                                                          PRT; 1075 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    298 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
  'PRINTS; PR00619; GATAZNFINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P17429; 4GAT.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                88.9%;
88.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GATA transcription factor.
                                                                                                                                                                           44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       43; Conservative
                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Penicillium chrysogenum.
                                                                                                                                              Best Local Similarity
Matches 44; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=5334;
                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         059842;
                                                                                                                                                                                                                                                                                                                                                                        Q9P952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  059842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 6
DR DR DR SO SO
                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                          Op
                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
```

```
AMELINE-21206224; PubMed-11309119;
AM MEDLINE-21206224; PubMed-11309119;
ACONION H., Zadra I., Haas H., Arst H.N. Jr., Jones M.G., Caddick M.X.;
AT The Aspergillus nidulans GATA transcription factor gene areB encodes
AT at least three proteins and features three classes of mutation.";
AL Mol. Microbiol. 40:361-375(2011).
BENBL; AF320976; AAG49353.1; -.
BREL; AF320976; AAG49353.1; -.
BREL; AF320976; AAG49353.1; -.
BREL; AF320976; AAGA9353.1; -.
BREL; AF320976; AAGA9353.1; -.
BREL; AF320976; AAGA9353.1; -.
BRELYES; PRO0619; GATA&INGER.
BRINTS; PRO0619; GATA&INGER.
BRENTE; PS001344; GATA_ZNFINGER.
BRENTE; PS001344; GATA_ZNFINGER.
BRENTE; PS001344; GATA_ZNFINGER.
SQ SEQUENCE 312 AA; 34279 MW; 35A73B7526FE5F88 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SECUENCE FROM N.A.
MEDLINE-89218944; PubMed-2651886;
ARDLINE-89218944; PubMed-2651886;
Arst H.N. Jr., Tollervey D., Caddick M.X.,
"A translocation associated, loss-of-function mutation in the nitrogen metabolite repression regulatory gene of Aspergillus nidulans can
                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                  6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Emericella nidulans (Aspergillus nidulans).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 312;
                                                                                                                                                                Length 298;
                                                                                                                                                                                                                                                                  50
                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                             17 CONCGTSTTPLWRRDELGSVLCNACGLFLKLHGRPRPISLKTDVIKSRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 CINCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR
PRINTS; PRO0619; GATAZNEINGER.
SMART; SMO0401; ZDE GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS550114; GATA_ZN_FINGER_2; 1.
SEQUENCE 298 AA; 32670 MW; C220D36F4993C590 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (TrEMBLrel. 16, Last sequence update) (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (TrEMBLrel. 16, Created)
(TrEMBLrel. 16, Last sequence update)
(TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                             10;
                                                                                                                                                             74.9%; Score 209; DB 3; 76.0%; Pred. No. 3.2e-20; tive 2; Mismatches 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 203; DB 3;
Pred. No. 2.2e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     312 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         320 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Emericella nidulans (Aspergillus nidulans).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        revert intracistronically."; Mol. Gen. Genet. 215:364-367(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           72.8%; Scur
74.0%; Pred
2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (TrEMBLrel. 16,
                                                                                                                                                                                        Best_Local Similarity 76.0
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AREB alpha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AREB beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAR-2001 (
01-MAR-2001 (
01-JUN-2002 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAR-2001 (
01-JUN-2002 (
GATA factor P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GATA factor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      37;
```

46251 MW; F91457AE2F8F7BB6 CRC64;

436 AA;

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9NJ18
Q9NJ18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q8WQQ4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08W004
                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                     09NJ18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q8WQQ4
   ÖS
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HDD BENEFIELD BE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                              QΥ
                                                                                                                                                                                                                                                                                                                                                                                                WEDLINE-21206224; PubMed=11309119;
WA MEDLINE-21206224; PubMed=11309119;
A Conlon H., Zadra I., Haas H., Arst H.N. Jr., Jones M.G., Caddick M.X.;
The Aspergillus nidulans GATA transcription factor gene areB encodes
T The Aspergillus nidulans GATA transcription factor gene areB encodes
T The Aspergillus nidulans GATA transcription factor gene areB encodes
T The Aspergillus nidulans and features three classes of mutation.";
Mol. Microbiol. 40:361-375(2001).
R FMBL; AR720976; AGAT.
R FMSE; P17429; 4GAT.
R FAGTO GATA. 1.
SMRT; SMO0409; GATAZNFINGER.
SMRT; SMO0401; ZnF_GATA.
R FRONTE; PSO0344; GATA_ZNFINGER_1; 1.
R PROSITE; PSO0344; GATA_ZNFINGER_2; 1.
SMRT; SMO0401; ZnF_GATA_ZNFINGER_2; 1.
SMRT; SMO0401; ZnF_GATA_ZNFINGER_2; 1.
SMRT; SMO0401; ZnF_GATA_ZNFINGER_2; 1.
SMRT; SMO0401; ZnF_GATA_ZNFINGER_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-89218944; PubMed-2651886;
Arst H.N. Jr., Tollervey D., Caddick M.X.;
A translocation associated, loss-of function mutation in the nitrogen metabolite repression regulatory gene of Aspergillus nidulans can revert intracistronically.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            $EQUENCE FROM N.A.

MEDLINE-21206224; PubMed-11309119;

MEDLINE-21206224; PubMed-11309119;

Conlon H., Zadra I., Haas H., Arst H.N. Jr., Jones M.G., Caddick M.X.;

The Aspergillus nidulans GATA transcription factor gene areB encodes
at least three proteins and features three classes of mutation.";

Mol. Microbiol. 40:361-375(2001).

EMBL, Microbiol. 46:351-1; -.
                                                                                                                                MEDLINE-89218944; PubMed-2651886; Arst H.N. Jr., Tollervey D., Caddick M.X.; Analevey D., Caddick M.X.; Atansiocation associated, loss-of-function mutation in the nitrogen metabolite repression regulatory gene of Aspergillus nidulans can
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
NCBL_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            72.8%; Score 203; DB 3; Length 320; 74.0%; Pred. No. 2.2e-19; tive 2; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 CINCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 CONCGTSKTPLWRRDELGSVLCNACGLFLKLHGRPRPISLKTDVIKSRNR 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (TrEMBLrel. 16, Created)
(TrEMBLrel. 16, Last sequence update)
(TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Emericella nidulans (Aspergillus nidulans).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
PRINTS; PR00619; GATAZNEINGER.
SMART; SM00401; ZnF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
   Eurotiales; Trichocomaceae; Emericella
                                                                                                                                                                                                                                                                          ct intracistronically.";
Gen. Genet. 215:364-367(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mol. Gen. Genet. 215:364-367(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GATA factor AREB gamma (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                   SEQUENCE FROM N.A.
                                NCBI_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001
01-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
                                                                                                                                                                                                                                                                              revert
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          O9HEV4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
Q9HEV4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AREB
OC OX REPRESENT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       අ
```

```
ö
                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wulbeck C., Simpson P.; "Expression of achaeter-scute homologues in discrete proneural clusters on the developing notum of the medfly Ceratitis capitata, suggests a common origin for the stereotyped bristle patterns of higher
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ceratitis capitata (Mediterranean fruit fly).

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota: Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Tephritoidea; Tephritidae; Ceratitis.

NCBI_TaxID=7213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Oestroidea; Calliphoridae; Calliphora.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                          Gaps
                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                        165 CTNCGTRTTTLWRRNNEGEPVCNACGLYFKLHGVNRPLAMRKDGIQTRKR 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                    11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wuelbeck C., Simpson P.;
Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AFIB4154; AAF61481.1; -.
HSSP: P17679; 10NF.
                                                                                                              1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; F1/b/y; Lugnr.
InterPro; IPR001164; hRIP_like.
InterPro; IPR001669; Znf_GATA.
Ffam; PF00320; GATA; 2.
PRINTS; PR00619; CATAZNINGER.
SMART; SM00401; ZnF_GATA; 2.
PROSITE; PS00344; GATA_ZN_FINGER_1; 2.
PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
SEQUENCE 324 AA; 34785 MW; CFCESF9D3CBD290B CRC64;
                                                                                                                                                                                                                                                                                                                                                                01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-WAR-2002 (TrEMBLrel. 20, Created)
01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               72.0%; Score 201; DB 5; 66.0%; Pred. No. 4.2e-19;
Score 203; DB 3;
Pred. No. 3e-19;
                                                    2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zinc finger DNA binding protein pannier.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=20171265; PubMed=10704387;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pannier-like protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Development 127:1411-1420(2000).
  72.8%;
74.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33; Conservative
                                                          37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. STRAIN-BENAKEO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-BENAKEO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Diptera.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PANNIER
```

Gaps

; 0

Length 486;

Score 198; DB 5; Pred. No. 1.6e-18;

71.0%; Sco... 66.0%; Pred. No. 1... 9; Mismatches

```
Drosophila melanogaster (Fruit fly).
Bukaryota, Metazoa, Arthropoda, Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera: Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophillade; Drosophila.
                                                                                                                                       SEQUENCE FROM N.A.
Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas W., Park S.,
Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
                                                                                                                                                                                                                                                                              Celniker S.;
Submitted (APR-2002) to the EMBL/GenBank/DDBJ datābases.
EMBL, AY095095; AAM1423.1; -
SEQUENCE 486 AA; 51606 MW; BCOE10E9BBD40A0B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search completed: January 3, 2003, 19:50:34
Job time : 49 secs
                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 66.0 les 33; Conservative
                                                                      Ephydroidea; Dros
NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09W6U0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             009W60
                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09W60
       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCOCCOS NO THE STANCE OF THE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wuelbeck C., Simpson P.;
Wuelbeck C., Simpson P.;
Wuelbeck C., Simpson P.;
William Sepression of pannier and achaete-scute homologs in a mosquito suggests an ancient role of pannier as a selector in the regulation of the dorsal body pattern.";
Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF395080; AAK97462.1;
InterPro; IPR00164; MIP_like.
InterPro; IPR00164; MIP_like.
Pfam; PF00320; GATA, 2.
PROSITE: PS00344; GATA_2N_FINGER_1; UNKNOWN_2.
PROSITE: PS50114; GATA_2N_FINGER_2; 2.
SEQUENCE 537 AA; 55906 MM; A25BCOCF53EF17B0 CRC64;
                                                                                              οĘ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 20, Last annocation update)
21-MAR-2002 (TrEMBLrel. 20, Last annocation update)
21nc finger transcription factor pannier.
Anopheles gambiae (African malaria mosquito).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Anopheles.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                      SEQUENCE FROM N.A.
Pistillo D., Skaer N., Simpson P.;
"Soute expression in Calliphora vicina reveals an ancestral pattern longitudinal stripes on the thorax of higher Diptera.";
Development 0:0-0(2002).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180 CINCGIRITILIRINEGEPVCNACGLYFKLHGVNRPLAMRKDGIQTRKR 229
                                                                                                                                                                                                                                                                                                                                                                                                                     Length 511;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                       511 AA; 53953 MW; 2E308EC61FEDFA33 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                72.0%; Score 201; DB 5; L 66.0%; Pred. No. 6.6e-19; ative 10; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ώ
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71.3%; Score 199; DB 5; 66.0%; Pred. No. 1.3e-18; tive 9; Mismatches 8;
                                                                                                                                                                                                                                                 PRINTS; PRO0619; GATAZNFINGER.
SMART; SMO0401; ZnF_GATA; 2.
PROSITE; PS00344; GATA_ZN_FINGER_1; UNKNOWN_2.
PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
NON_TER 511 511 511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                537 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       486 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                     EMBL; AY061878; AAL32068.1; -.
InterPro; IPR0001164; hRIP_like.
InterPro; IPR000159; Znf_GATA.
Pfam; PF00320; GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21,
21,
21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 66.0
Matches 33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-JUN-2002 (TrEMBLrel.
01-JUN-2002 (TrEMBLrel.
01-JUN-2002 (TrEMBLrel.
SD02611p.
                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 33: Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=7165;
NCBI_TaxID=7373;
                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q8T3J0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q95VY5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
Q95VY5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q813J0
ID Q8
AC Q8
DT 01
DT 01
DE SD
GN PN
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-EMBRYO;
MEDLINE-99307072; PubMed-10375499;
REDLINE-99307072; PubMed-10375499;
Smith J.C., Patient R., Holder N.;
Induction of the mesendoderm in the zebrafish germ ring by yolk cell derived TGF-(beta) family signals and discrimination of mesoderm and endoderm by FGF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                    Brachydanio rerio (Zebrafish) (Zebra danio).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 383;
1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69.9%; Score 195; DB 13; Lengtn 5:
64.0%; Pred. No. 3.2e-18;
64.0%; Pred. No. 3.2e-18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AST, 2DB-GENE-98026-340; gata5.
InterPro; IPR001164; hRIP_like.
InterPro; IPR001164; hRIP_like.
InterPro; IPR001164; hRIP_like.
Fram; PF00320; GATA2.
PRINTS; PR00619; GATAZNFINGER.
SWART; SM00401; ZnF_GATA; 2.
PROSITE; PS00144; GATA_ZN_FINGER_1; 2.
PROSITE; PS01144; GATA_ZN_FINGER_2; 1.
SEQUENCE 383 AA; 41479 MW; 7F752664185588AF CRC64;
                                                                                                                                                                                                                         01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
Gata5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Development 126:3067-3077(1999).
EMBL; AJ242515; CAB43400.1; -.
HSSP; P17679; 1GNF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   32; Conservative
                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=7955
                                                                                                                                                                                                                                                                                                                                               GATAS OR GTAS.
```

```
January 3, 2003, 18:50:12; Search time 14 Seconds (without alignments) 148.130 Million cell updates/sec
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                              OM protein - protein search, using sw model
```

Run on:

1 CINCFIQITPLWRRNPDGQP.....LHGVVRPLSLKTDVIKKRNR US-09-725-010-3 score: Perfect

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table: Sednence:

Notal number of hits satisfying chosen parameters:

112892 segs, 41476328 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Listing first 45 summaries Post-processing: Minimum Match 0% Maximum Match 100%

SwissProt\_40:\* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

|            | Description  |            |            |            |            |            |            |            |            | _          |            | P43574 saccharomyc | -          |            |            |            | P43696 xenopus lae |            | _            | -          |           | Q91678 xenopus lae | -          |            | gallus g   | xenobns    |            |          | _          |            |            | P43691 gallus gall | snw        | Q08369 mus musculu |
|------------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------------|------------|------------|------------|------------|--------------------|------------|--------------|------------|-----------|--------------------|------------|------------|------------|------------|------------|----------|------------|------------|------------|--------------------|------------|--------------------|
|            | QI           | NIT2_NEUCR | AREA_PENCH | AREA_PENRO | NRFA_PENUR | AREA_ASPOR | AREA_EMENI | AREA_ASPNG | AREA_GIBFU | NUT1_MAGGR | GAF1_SCHPO | GAT1_YEAST         | ASD4_NEUCR | GZF3_YEAST | DA80_YEAST | GLN3_YEAST | GA5B_XENLA         | GA5A_XENLA | PNR_DROME    | GAT5_MOUSE | SRP_DROME | GA6A_XENLA         | GAT5_CHICK | GAT5_HUMAN | GAT6_CHICK | GA6B_XENLA | GAT4_XENLA | GAT6_RAT | GAT6_MOUSE | GAT6_HUMAN | URB1_USTMA | GAT4_CHICK         | GAT1_MOUSE | GAT4_MOUSE         |
|            | B            | 1          | 1          | ٦          | ٦          | П          | ٦          | Н          | -1         | Н          | -          | -                  | Н          | П          | ٦          | П          | <del>, ,</del>     | Н          | <del>,</del> | Н          | ٦         | Н                  | Н          | ~          | Н          | -          | Н          | ~        | -          | 7          | Н          | 7                  | Н          | -                  |
|            | Match Length | 1036       | 725        | 860        | 865        | 866        | 876        | 882        | 971        | 926        | 855        | 510                | 426        | 551        | 269        | 730        | 388                | 390        | 540          | 404        | 779       | 391                | 391        | 397        | 387        | 391        | 392        | 441      | 444        | 449        | S          | 380                | 413        | 440                |
| *<br>Query | Match        | 100.0      | 98.6       | 98.6       | 98.6       | 98.6       | 98.6       | 98.6       | 98.6       | 95.7       | 89.7       | 82.2               | 75.1       | 74.7       | 72.6       | 71.5       | 69.4               | 4.69       | 68.7         | 67.6       | 9.79      | 67.3               | 67.3       | 67.3       | 6.99       | 6.99       | 66.99      | 6.99     | 6.99       | 6.99       | 66.5       | 65.8               | 65.1       | 65.1               |
|            | Score        | 281        | 277        | 277        | 277        | 277        | 277        | 277        | 277        | 269        | 252        | 231                | 211        | 210        | 204        | 201        | 195                | 195        | 193          | 190        | 190       | 189                | 189        | 189        | 188        | 188        | 188        | 188      | 188        | 188        | 187        | 185                | 183        | 183                |
| Result     | No.          | 1          | 7          | ю          | 4          | S          | 9          | 7          | σο         | 6          | 10         | 11                 | 12         | 13         | 14         | 15         | 16                 | 17         | 18           | 19         | 20        | 21                 | . 22       | 23         | 24         | 25         | 56         | 27       | 28         | 29         | 30         | 31                 | 32         | 33                 |

| P46152 rattus norv<br>P43694 homo sapien | P23767 xenopus lae P15976 homo sapien P43479 rattus porv | P28515 caenorhabdi<br>P23768 xenopus lae | Q10134 schizosacch<br>P23770 xenopus lae | P91623 drosophila<br>P17678 gallus gall | P23824 gallus gall |
|--|--|--|--|---|--------------------|
| GAT4_RAT<br>GAT4_HUMAN                   | GA1A_XENLA<br>GAT1_HUMAN                                 | GA11_CAEEL<br>GA1B_XENLA                 | GAF2_SCHPO<br>GAT2_XENLA                 | GATC_DROME<br>GAT1_CHICK                | GAT2_CHICK         |
| нг                                       | 44   |  | н н                                      |   | H                  |
| 440                                      | 359<br>413   | 416<br>364                               | 564<br>452                               | 486<br>304                              | 466                |
| 65.1<br>65.1                             | 64.8<br>8.8<br>8.8                                       | 64.8<br>64.8                             | 64.4<br>64.1                             | 64.1                                    | 63.7               |
| 183                                      | 182  | 182<br>181                               | 181                                      | 180<br>179                              | 179                |
| 9.84<br>5.6                              | 37   | 2 8 4<br>2 0 0                           | 41                                       | 43<br>44                                | 45                 |

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                MEDITINE 901588568; PubMed=2137552; Fu Y.-H., Marzluf G.A.; introgen regulatory gene of Neurospora crassa, "nit.2, the major nitrogen regulatory gene of Neurospora crassa, encodes a protein with a putative zinc finger DNA-binding domain."; Mol. Cell. Biol. 10:1056-1065(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRINTS; PRO0619; GATAZNFINGER.
SMART; SMO0401; ZnF_GATA; 1.
PROSITE; PSO0344; GATA_ZN_FINGER_1; 1.
PROSITE; PSS0114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordarlomycetes;
Sordariales; Sordariaceae; Neurospora.
NCBI_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -!- SUBCELLULAR LOCATION: Nuclear.
-!- INDUCTION: BY LACK OF A PRIMARY NITROGEN SOURCE.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                          (Rel. 16, Last sequence update) (Rel. 40, Last annotation update)
                                                                                                        Nitrogen catabolic enzyme regulatory protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSFAC; T00627; -.
Interpro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                            01-NOV-1990 (Rel. 16, Created)
01-NOV-1990 (Rel. 16, Last seq
16-OCT-2001 (Rel. 40, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M33956; AAB03891.1; -.
                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A34755; A34755.
                                                                                                                                                                                                                                 STRAIN=74-OR23-1A;
                                                                                                                                         Neurospora crassa.
                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                         [2]
MUTAGENESIS.
                                 NITZ_NEUCR
RESULT 1
NIT2_NEUCR
```

525 CINCFIQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 574

```
.;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE=95308537; PubMed=7788718;
Haas H., Bader B., Redl B., Stoeffler G., Marzluf G.A.;
Haber and analysis of nre, the major nitrogen regulatory
gene of Penicillium chrysogenum.";
Curr. Genet. 27:150-158(1995).
                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bukaryota; Fungi; Ascomycota; Pezizomycotina; Burotiomycetes;
Burotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium
                                                                                       CTNC->STNG: ABOLISHES DNA-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                         Length 1036;
                                                                                                                                GL->DV: ABOLISHES DNA-BINDING.
RR->NS: ABOLISHES DNA-BINDING.
109310 MW; 5FE4992BIC223514 CRC64;
                                                                                                                                                                                                                                                              RR->GG: ABOLISHES DNA-BINDING.
                                                                                                                  NA->DV: ABOLISHES DNA-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98.6%; Score 277; DB 1; Length 725; 98.0%; Pred. No. 5.7e-28; Live 1; Mismatches 0; Indels
                                                                                                                                                                                                                                               1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-JNN-2002 (Rel. 41, Last annotation update)
Nitrogen regulatory protein areA (Nitrogen regulator nre).
                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45C24148F81F8D43 CRC64;
                3 X APPROXIMATE REPEATS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1- SUBCELLULAR LOCATION: Nuclear.
-1- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                      100.0%; Score 281; DB 1; 100.0%; Pred. No. 2.4e-28;
Nuclear protein; Nitrate assimilation; Repeat.
DOMAIN 49 110 3 X APPROXIMATE R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98.0%; Pred. ac.
                                                                                                                                                                                                                    0; Mismatches
                                                                         GATA-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nuclear protein; Nitrate assimilation.
2N FING 525 549 GATA-TVDF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Princer; Proof Sara; 1.
PRINTS; PROO619; GATAZNFINGER.
SMART; SMOO401; ZAF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             725 AA; 76848 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U02612; AAA83400.1; -.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      49; Conservative
                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Penicillium chrysogenum.
                            55
92
1110
767
746
756
756
                            49
105
143
743
755
755
765
768
768
768
769
1036 AA;
                                                                                                                                                                                                     Similarity
50; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=5076;
                                                                                                                                                                                                                                                                                                                                                                                                                                   AREA OR NRE
                                                                                                                                                                                                                                                                                                                                              AREA_PENCH
Q01582;
                                        REPEAT
REPEAT
ZN_FING
MUTAGEN
                                                                                                                                                           SEQUENCE
                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                  MUTAGEN
MUTAGEN
                                                                                                                              MUTAGEN
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                       Local
                              REPEAT
                                                                                                                                                                                                                                                                                                                                    AREA_PENCH
                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                     RESULT 2
KY
FT
FT
FT
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                       g
```

1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50

ŏ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                        Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger; Nuclear protein; Nitrate assimilation.
ZN_FING 660 660 67AA-TYPE.
SEQUENCE 860 AA; 91615 MW; ZB8BF0DDCDC1BA68 CPC64.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 860;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          660 CINCFIQITPLMRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 709
                                                                                                                      Nitrogen regulatory protein area (Nitrogen regulator nmc).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98.6%; Score 277; DB 1; Length 86 98.0%; Pred. No. 6.7e-28; tive 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                            Gente S., Poussereau N., Fevre M.;
Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN.
                                                                                                                                                                                                                                                                                                                               -!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                               15-JUL-1998 (Rel. 36, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ellis C.M.;
Thesis (1996), University of Calgary, Canada.
                                  860 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             865 AA
                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nitrogen regulatory protein NRFA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AJ001530; CAA04815.1; -. HSSP; P17429; 4GAT. InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PP00320; GATA; 1.
PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91615 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   49; Conservative
                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                        Penicillium roqueforti.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Penicillium urticae.
                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-NRRL 2159A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                         NCBI_TaxID=5082;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=29844;
                                                                                                                                        AREA OR NMC
                                  AREA PENRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NRFA_PENUR
Q92269;
                                                                                                                                                                                                                                                              STRAIN-P2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
RESULT 3
AREA_PENRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NRFA_PENUR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
```

ö

us-09-725-010-3.rsp

```
repression.
                                                                                                                                                                                                                                                                                                                                                                                      AREA_EMENI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REVISIONS.
                                                              ZN_FING
SEQUENCE
                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                              AREA_EMENI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB;
                                                                                                                                                                                                                                                                                                                                                                                                              NA KW
KW
SO
SO
                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions and also so its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL OUTSTAILD. the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-IFO 4177;
Christensen T., Hynes M.J., Davis M.A.;
Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN. POSITIVELY ACTING
REGULATORY GENE OF NITROGEN METABOLLIE REPRESSION (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PRO0619; CARAZNEINGER.
SMART; SM00401; ZAFAZNEINGER.
PROSITE; PS00344; GATA_ZN_FINGER.1; 1.
PROSITE; PS50114; GATA_ZN_FINGER.2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
Nuclear profefi; Nitrate assimilation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 865;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   665 CINCFTQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98.6%; Score 277; DB 1; Length 86 98.0%; Pred. No. 6.7e-28; Live 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GATA-TYPE. ; CC901BED187D8EEF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1- SUBCELLULAR LOCATION: Nuclear.
-1- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                          -!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
  FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PROD619; GATAZNFINGER.
SMART; SMO0401; ZNF_GATA; 1.
PROSITE; PSO0344; GATA_ZN_FINGER_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nitrogen regulatory protein areA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AJ002968; CAA05776.1; -.
                                                                                                                                                                                                                                                                                                                     TRANSFAC; T02828; -.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       865 AA; 92407 MW;
                                                                                                                                                                                                                                                                        EMBL; U53137; AAB17740.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00320; GATA; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Aspergillus oryzae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=5062;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AREA_ASPOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZN_FING
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               013415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
AREA_ASPOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SO DE RESERVA DE RESER
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE-90228331; PubMed-1970293; Kudla B., Caddick M.X., Langdon T., Martinez-Rossi N.M., Kudla B., Caddick M.X., Langdon T., Martinez-Rossi N.M., Bennett C.F., Sibley S., Davies R.W., Arst H.N. Jr.; "The regulatory gene area mediating nitrogen metabolite repression in Aspergillus nidulans. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger."; EMBO J. 9:1355-1364(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Langdon T., Seerins A., Ravagnani A., Caddick M.X., Arst H.N. Jr., "Mutational analysis reveals dispensability of the N-terminal region of the Aspergillus transcription factor mediating nitrogen metabolite
                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         factors.";
J. Mol. Biol. 277:605-620(1998).
-!- FUNCTION: TRANSCRIPTION ACTIVATOR THAT MEDIATES NITROGEN
METABOLITE REPRESSION IN A.NIDULANS. ACTIVATE THE TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "The solution structure of a fungal AREA protein-DNA complex: an alternative binding mode for the basic carboxyl tail of GATA
                                                                                                                                                                                                                                            ö
                Transcription regulation, Activator; DNA-binding; Zinc-finger; Nuclear protein; Nitrate assimilation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Emericella nidulans (Aspergillus nidulans).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
                                                                                                                                                                                                                                                                                                                                         Length 866;
                                                                                                                                                                                                                                                                                                        1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Starich M.R., Wikstroem M., Arst H.N. Jr., Clore G.M., Gronenborn A.M.;
                                                                                                                 C09A03EB12E3FBB4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -:- SUBCELLULAR LOCATION: Nuclear.
-:- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                        ..
                                                                                                                                                                                98.6%; Score 277; DB 1; 98.0%; Pred. No. 6.7e-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-FEB-1996 (Rel. 33, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   876 AA
                                                                                                                                                                                                          Pred. No. 6.7e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Fungi, Ascomycota; Pezizomyc
Eurotiales; Trichocomaceae; Emericella.
NCBL_TaxID=5072;
                                                                                      GATA-TYPE
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=98202574; PubMed=9533883;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=96123430; PubMed=8596437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nitrogen regulatory protein areA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mol. Microbiol. 17:877-888(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1990 (Rel. 15, Created)
                                                                                                                 92882 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; X52491; CAA36731.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [3]
STRUCTURE BY NMR OF 662-727.
                                                                                                                                                                                                              Best Local Similarity 98.09
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4GAT; 28-JAN-98.
5GAT; 28-JAN-98.
6GAT; 28-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S10017
                                                                                                                    866 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S10017;
```

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Maccabe A.P., Vanhanen S.A.S., Sollewijn Gelpke M.,
van de Vondervoort P., Arst H.N., Visser J.;
Submitted (SEP-194) to the EMEL/GenBank/DBJ databases.
-!- FUNCTION: MAJOR NITRGEN REGULATORY PROTEIN. POSITIVELY ACTING
REGULATORY GENE OF NITROGEN METABOLITE REPRESSION.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- SUBCELLULAR CORTAINS 1 GATA-TYPE ZINC FINGER.
                                           INCELTUS, IEANOUSON, I.
PERMI, PE00320; GATAZNEINGER.
SMART, SM00401; ZnF_GATAZ, 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS501144; GATA_ZN_FINGER_2; 1.
PROSITE; PS501145; GATA_ZN_FINGER_2; 1.
TRANSCTIPTION regulation; Activator; DNA-binding; Zinc-finger; Nuclear protein; Nitrate assimilation; 3D-structure.
ZN_FING GATA TYPE.
H-T-H MOTIF (PROBABLE).
                                                                                                                                                                                                                                                                                              0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                     98.6%; Score 277; DB 1; Length 876; llarity 98.0%; Pred. No. 6.8e-28; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                         Query Match 98.6%; Score 277; DB 1; Length 882; Best Local Similarity 98.0%; Pred. No. 6.8e-28; Matches 49; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                          1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                 721 742 H-T-H MOTIF (PROBABLE).
876 AA; 94195 MW; 9ADC2273EE536F98 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            676 700 GATA-TYPE.
882 AA; 94518 MW; 97A502936B94E5AE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Rel. 36, Created)
(Rel. 36, Last sequence update)
(Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nuclear protein; Nitrate assimilation.
ZN_FING 676 700 GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Nitrogen regulatory protein areA.
PDB; 7GAT; 28-JAN-98.
TRANSFAC; T02533; -.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR00069; Znf_GATA.
Pfam; PF00320; GATA; 1.
PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; X81998; CAA57524.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Aspergillus niger.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P17429; 4GAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=5061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-N402;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-JUL-1998
15-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16-OCT-2001
                                                                                                                                                                                                                                                                                           49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  AREA_ASPNG
                                                                                                                                                                                                 DNA_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                   Tudzynski B., Homann V., Feng B., Marzluf G.A.;
"Isolation, characterization and disruption of the area nitrogen
regulatory gene of Gibberella (dijkuroi.";
Mol. Gen. Genet. 261:106-114(1999).
-!- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN. POSITIVELY ACTING
REGULATORY GENE OF NITROGEN METABOLITE REPRESSION (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                         Gibberella fujikuroi (Bakanae and foot rot disease fungus) (Fusarium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                           Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
Nuclear protein; Nitrate assimilation.
ZN_FING 694 718 GATA_TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ó;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).
Eukaryota; Pungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
98.6%; Score 277; DB 1; Length 971; 98.0%; Pred. No. 7.5e-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    971 AA; 103580 MW; 887DD882141C7453 CRC64;
                                                                                                                                           15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             956 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98.0%; Pred. No. 7.5e
Live 1; Mismatches
                                                                                                                                                                                                                                                                                              Hypocreales; Nectriaceae; Gibberella,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=99168774; PubMed=10071216;
                                                                                                                                                                                                      Nitrogen regulatory protein areA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nitrogen regulatory protein NUT1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
PRINIS; PR00619; GATANFINGER.
SMART; SMO0401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; Y11006; CAA71897.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             49; Conservative
                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                             NCBI_TaxID=5127;
                                                                                                                                                                                                                                                          moniliforme)
                                                                                                          AREA_GIBFU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NUT1_MAGGR
Q01168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                           AREA_GIBFU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NUT1_MAGGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NOT1.
                                                                        RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 9
                   g
                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AC DIT DIT OCC OCC OCC OCC
```

ó

1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50

ð

Mon Jan

```
δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mach by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mood V., Gaillam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Goble A., Hamin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Ollyer K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 663 CINCATQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 712
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 956;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                       95.7%; Score 269; DB 1; Length 95r
96.0%; Pred. No. 7.9e-27;
' Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                           40ABDA5A07A7D7AB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Transcription factor gafl (Gaf-1).
GAFI OR SPCGA17.01C OR SPCG1902.01.
Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                     -i- SUBCELLULAR LOCATION: Nuclear.
-i- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      010280; 094482; 09USK9; 01-NOV-1997 (Rel. 35, Created) (COTT-2001 (Rel. 40, Last sequence update) 15-JUN-2002 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 855 AA.
                                                                                                                                                                                                                                                                                                                                                                                                Nuclear protein; Nitrate assimilation.
ZN_FING 663 687 GATA-TYPE
SEQUENCE 956 AA; 100874 MW: 40APDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=21848401; PubMed=11859360;
                                       STRAIN=Guyane 11;
MEDLINE=96335139; PubMed=8757395;
                                                                                                                                                                                                                                                                                                 TRANSFAC; T02829; -.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                      PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZNF_GATA; 1.
                                                                                                                                                                                                                                                                          EMBL; U60290; AAB03415.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 96.0% nes 48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Schizosaccharomyces.
NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GAF1_SCHPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

```
this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its modified and this statement is so not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
A Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Noodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
A Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
A Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
B Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
B Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
A Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
A Golfeau A., Rochet M., Gaillardin C., Moore K., Hurst S.M.,
A Daga R.R., Cruzado L., Jimene Z., Tallada V.A., Galzon A., Thode G.,
A Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
The genome sequence of Schizosaccharomyces pombe.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE OF 566-855 FROM N.A. MEDLINE=98382525; PubMed=9714831; Hoe K.-L., Won M.S., Chung K.-S., Park S.-K., Kim D.-U., Jang Y.-J., Yoo O.-J., Yoo H.-S.; "Molecular cloning of gafl, a Schizosaccharomyces pombe GATA factor, which can function as a transcriptional activator.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetales
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Interprof. 17800679; Znf_GATA.

Pfam; PF00320; GATA; 1.

PRINTS; PR00619; GATAZNFINGER.

SMART; SM00401; ZnF_GATA. 1.

PROSITE; PS00344; GATA_ZN_FINGER. 1; 1.

PROSITE; PS50114; GATA_ZN_FINGER_2; 1.

Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 89.7%; Score 252; DB 1; Length 855; Best Local Similarity 88.0%; Pred. No. 1.1e-24; Matches 44; Conservative 5; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 252; cc.,
Pred. No. 1.16-24;
Pred. Transhes 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            855 AA; 91776 MW; 3D932F83D2DE6774 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -i - SUBCELLULAR LOCATION: Nuclear (Probable).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GAT1 OR YFL021W.
Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Transcriptional regulatory protein GAT1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gene 215:319-328(1998).
-!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AL035076; CAA22647.1; -.
EMBL; AL04951; CAB40003.1; -.
EMBL; L31601; AAC35593.1; -.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSFAC; T02831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAT1_YEAST
P43574;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 11
```

ö

```
Matches
                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ID DAT DOT DATE OF THE PROPERTY OF THE PROPERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseeisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                  MEDLINE-96182087; PubMed-8622686; Coffman J.A., Rai R., Cunningham T., Svetlov V., Cooper T.G.; Gatlp, a GATA family protein whose production is sensitive to nitrogen catabolite repression, participates in transcriptional activation of nitrogen-catabolic genes in Saccharomyces cerevisiae."; Mol. Cell. Biol. 16:847-858(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- FUNCTION: POSTITYE REGUÍATOR OF MULTIPLE NITROGEN CATABOLIC GENES.
-!- SUBCELLULAR LOCATION: Nuclear (Potential).
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRINTS; PRO0619; GARAZINEINGER.
SMART; SM00401; ZnF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
FRANSCription regulation; DNA binding; Zinc-finger; Nuclear protein.
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                         Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M., Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K., Yamazaki M., Tashiro H., Eki T.; Analysis of the nucleotide sequence of chromosome VI from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-OCT-2001 (Rel. 40, Last annotation update)
GATA type zinc finger protein asd4 (Ascus development protein 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLY-ASP.
: 62D805E42695F35F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 231; DB 1;
Pred. No. 3.3e-22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Rel. 40, Last sequence update) (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      426 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A., AND CHARACTERIZATION. MEDLINE-20455773; PubMed=10998244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82.2%; Score 231;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN-S288c / AB972;
MEDLINE-95400292; PubMed-7670463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGD; S0001873; GAT1.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16-OCT-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             510 AA; 56327 MW;
                                                                                                                                                                                                                                                                                                                                                                                         Saccharomyces cerevisiae.";
Nat. Genet. 10:261-268(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; U27344; AAB03516.1; -.
EMBL; D50617; BAA09217.1; -.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               80.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSFAC; T02817;
  NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ASD4_NEUCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9HEV5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ASD4_NEUCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               datches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                          BIOCOCHEMISELY 39:11065-11073(2000).

-!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT FUNCTIONS IN SEXUAL DEVELOPMENT; DISRUPTION OF ASD-4 GENE RESULTS IN AGENESIS OF ASCUS AND ASCOSPORE WITH MACROSCOPICALLY NORMAL FRUITING BODY FORMATION. THE GATA-TYPE ZINC FINGER DOMAIN BINDS TO DNA SEQUENCES FROM ITS OWN PROMOTER REGION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Rasmussen S.W.;
"A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and CSD3 genes, a TCP-1-related gene, an open reading frame similar to the DAL80 gene, and a tRNA(Arg).";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS; PRO0619; GATAZNFINGER.
SMART; SMO401; ZDF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; DNA-binding; Zinc-finger; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
Feng B., Haas H., Marzluf G.A.; "ASD4, a new GATA factor of Neurospora crassa, displays sequence-specific DNA binding and functions in ascus and ascospore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=S288C / FY1679;
MEDLINE=97103775; PubMed=8948101;
Cziepluch C., Kordes E., Pujol A., Jauniaux J.-C.;
"Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 211; DB 1; Length 426;
Pred. No. 1e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16 CQNCATSTTPLWRRDEMGQVLCNACGLFLKLHGRPRPISLKTDVIKSRNR 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COILED COIL (POTENTIAL).
4B7DE193050E280E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Saccharomycetales; Saccharomycetaceae; Saccharomyces. NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                       SUBUNIT: HOMOTETRAMER.
SUBCELLULAR LOCATION: Nuclear.
SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 551 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. le-1
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=S288c / FY1679;
MEDLINE=96090136; PubMed=7483851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; P17429; 4GAT.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; AF319953; AAG45180.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               292 Cr
46532 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Rel. 32, Created)
(Rel. 32, Last sequ
(Rel. 40, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            75.18;
78.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GZF3 protein.
GZF3 OR YJL110C OR J0806.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    39; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Yeast 11:873-883(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16
182 2
426 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEOUENCE FROM N.A.
                                                                                            development.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Coiled coil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GZF3_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Gutstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Expression of the DAL80 gene, whose product is homologous to the GATA factors and is a negative regulator of multiple nitrogen catabolic genes in Saccharomyces cerevisiae, is sensitive to nitrogen
                                                                                                                                                                                                                                                                                                                                                                                                                        FIGHT. FEOGLS: GATAZNEINGER.
SMART; SMO0401; ZnF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
TRANSCIPTION regulation; DNA binding; Zinc-finger; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,
             SPT10, GCD14, RPE1, PHÓ86, NCA3, ASF1, CCT7, G2F3, two tRNA genes, three remnant delta elements and a Ty4 transposon."; Yeast 12:1471-1474(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes,
Saccharomycetales, Saccharomycetaceae, Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Coornaert D., Vissers S., Andre B., Grenson M.; "The UGA43 negative regulatory gene of Saccharomyces cerevisiae contains both a GATA-1 type zinc finger and a putative leucine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-1992 (Rel. 22, Last sequence update)
16-COT-2001 (Rel. 40, Last annotation update)
Nitrogen regulatory protein DAL80 (Regulatory protein UGA43).
DAL80 OR UGA43 OR YRR034W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              131 CKNCLTSTTPLWRRDEHGAMLCNACGLFLKLHGKPRPISLKTDVIKSRNR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GATA-TYPE.
807243799BFAB704 CRC64;
                                                                        -1- SUBCELLULAR LOCATION: Nuclear (Potential).
-1- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                74.7%; Score 210; DB 1; 76.0%; Pred. No. 1.8e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cunningham T.S., Cooper T.G.;
Mol. Cell. Biol. 12:2454-2454(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    catabolite repression.";
Mol. Cell. Biol. 11:6205-6215(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-92049352; PubMed=1944286;
Cunningham T.S., Cooper T.G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE=92405226; PubMed=1525858;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=92236621; PubMed=1569960;
                                                                                                                                                                                                                                                                                                                                                                   SGD; S0003646; GZF3.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-1992 (Rel. 22, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             551 AA; 60519 MW;
                                                                                                                                                                                                                                                                           EMBL; X85021; CAA59384.1; -. EMBL; Z49385; CAA89405.1; -. EMBL; X86323; CAA60126.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity 76.09
nes 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00320; GATA; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DA80_YEAST
P26343;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DA80_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way mon-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDILINE-29049333; PubMed-1682800;
MEDILINE-29049333; PubMed-1682800;
MINICHINE-29049333; PubMed-1682800;
"Sequence and expression of GLN3, a positive nitrogen regulatory gene of Sarcharomyces cerevisiae encoding a protein with a putative zinc finger DNA-binding domain.";
MOI. Cell. Biol. 11:6216-6228(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces. NCBL_TaxID=4932;
                                                         Urrestarazu L.A., Jauniaux J.-C.;
Submitted (MAR-11994) to The EmBL/GenBank/DDBJ databases.
-!- FUNCTION: NEGATIVE REULATOR OF MULTIPLE NITROGEN CATABOLIC
GENES INCLUDING THE ALLANTOIN PATHWAY GENES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00401; ZnF_GATA; 1.
PROSTET; PRO0344; GATA_ZNLFINCER_1; 1.
PROSITE; PS50114; GATA_ZNLFINGER_2; 1.
Transcription regulation; Repressor; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 204; DB 1; Length 269;
                                                                                                                      -:- SUBCELLULAR LOCATION: Nuclear.
-:- INDUCTION: SENSITIVE TO NITROGEN CATABOLITE REPRESSION.
-:- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F37BEB10038599EC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ARG/LYS-RICH (BASIC).
S -> L (IN REF. 3).
V -> I (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.4e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1990 (Rel. 16, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GATA-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nuclear protein; Nitrate assimilation.
ZN_FING 31 55 GATA-TYPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ASN-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nitrogen regulatory protein GLN3. GLN3 OR YER040W.
                                                                                                                                                                                                                                                                                                         EMBL; M77821; AAA34556.1; --
EMBL; X60199; CAA42757.1; --
EMBL; Z28259; CAA82107.1; --
EMBL; Z28258; CAA82106.1; --
PIR; S22781; S22781.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRINTS; PR00619; GATAZNFINGER.
                                                                                                                                                                                                                                                                                                                                                                                                       TRANSFAC; T02411; -.
SGD; S0001742; DALB0.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA: 1.
           Curr. Genet. 21:301-307(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30166 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           72.68;
74.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 74.08
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55
100
108
6
207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31
79
101
6
207
269 AA;
                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLN3_YEAST
P18494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
zipper.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLN3_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

STAIN-S288c / AB972;

A Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,

A Alles E., Berno A., Emenan T., Carpenter J., Chen E., Cherry J.M.,

RA Annan R., Kayser A., Komp C., Lashkarl D., Lew H., Lin D.,

RA Hyman R., Kayser A., Namath A., Norgren R., Oefner P., Oh C.,

RA Mosedale D., Nakahara K., Namath A., Shogren T., Smith V.,

RA Apjor P., Wel Y., Yelcon M., Botstein D., Davis R.W.;

Raylor P., Wel Y., Yelcon M., Botstein D., Davis R.W.;

Rubmitted (DEC-1994) to the EmbL/Genbank/DDBJ databases.

-I- FUNCTION: POSITIVE NITROGEN REGULANORY PROTEIN. REQUIRED FOR

THE ACTIVATION OF TRANSCRIPTION OF A NUMBER OF GENES (INCLUDING

THE ALLANTOIN PATHWAY GENES) IN RESPONSE TO THE RELACEMENT OF

GLUTAMINE BY GLUTAMATE AS SOURCE OF NITROGEN. BINDS THE NITROGEN

UBSTREAM ACTIVATION SEQUENCE OF GLNI, THE GENE ENCODING GLUTAMINE

C. SYNTHETASE. UREZ MAY CAPALITICALLY INACTIVATE GLN3 IN RESPONSE TO

AN INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.

C. SIGLIAM INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.

C. SIGLIAM INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.

C. SIGLIAM INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.

C. SIGLIAM INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.

C. SIGLIAM INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.

C. SIGLIAM INCREASE IN THE INTRACELLULAR CONCENTRATION OF GLUTAMINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transcription regulation; Activator; DNA-binding; Zinc-finger;
Nuclear protein; Nitrate assimilation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71.5%; Score 201; DB 1; Length 730; 75.0%; Pred. No. 3.3e-18; Live 4; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              474 474 P -> G (IN REF. 1).
730 AA; 79382 MW; 3159E1844469942E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ARG/LYS-RICH (BASIC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRINTS; PRO0619; GATAZNFINGER.
SMARY; SMO401; ZnF_GATA; 1.
PROSTITE; PSS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SGD; S0000842; GLN3.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M35267; AAA34645.1; -. EMBL; U18796; AAB64575.1; -. PIR; S22280; S22280. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSFAC; T02818; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
```

Search completed: January 3, 2003, 19:48:37 Job time : 14 secs

```
GenCore version 5.1.3 Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

Run on:

January 3, 2003, 19:46:47; Search time 25 Seconds (without alignments) 192.269 Million cell updates/sec

US-09-725-010-3

281 1 CTNCFTQTTPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283224 segs, 96134422 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_73:\* Database

1: pirl:\* 2: pir2:\* 3: pir3:\* 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

major nitrogen regulation protein - Penicillium chrysogenum (strain Q176)
C.Species: Penicillium chrysogenum
C.Species: Penicillium chrysogenum
C.Species: Strain Q176
C.Sate: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change O8-Dec-2000
C.Sate: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change O8-Dec-2000
C.Sate: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change O8-Dec-2000
C.Sate: Strain Strai

δλ g RESULT 2 S51493

C;Superfamily: nitrogen regulatory protein nit-2; GATA-type zinc finger homology C;Keywords: zinc finger C;Keywords: zinc finger F;659-712/Domain: GATA-type zinc finger homology <GZF>

C;Genetics: A;Introns: 123/2

| 00000000000000   | A56953<br>T19677<br>ALIGNMENTS | it-2 - Neurospora crassa revision 26-Jul-1996 #te- 5, 1990 rogen regulatory gene of UID:90158568; PMID:213755 atory protein nit-2; GATA scription regulation; zin inc finger homology <gze>GATA motif</gze>   | 100.0%; Score 281; DB 1; Length 1036;<br>100.0%; Pred. No. 1e-27;<br>ive 0; Mismatches 0; Indels 0; Gaps 0; |
|--|--------------------------------|---|---|
| 179 63.7<br>179 63.7<br>179 63.7<br>179 63.7<br>179 63.7<br>179 63.7<br>178 63.3<br>178 63.3<br>178 63.3<br>175 62.3<br>175 62.3<br>171 60.9 |                                | RESULT 1 A34755  nitrogen regulatory protein nit-2 - C; Species: Neurospora crassa C; Date: 13-Uul-1990 #sequence_revisi C; Accession: A34755 R; Fu, Y. H.; Marzluf, G.A. Mol. Cell. Biol. 10, 1056-1065, 1990 A; Title: nit-2, the major nitrogen r A; Reference number: A34755; MUID:901 A; Accession: A34755 A; Molecule type: DNA; mRNA A; Residues: 11036 < CFUY> A; Cross-references: GB:M33956 C; Genetics: 209/2; 335/3 C; Superfamily: nitrogen regulatory p C; Reywords: DNA binding; transcriptic F;740-793/Domain: GATA-type zinc fing F;743-767/Region: zinc finger GATA m | Query Match Best Local Similarity 100 Matches 50; Conservative  |

Mon Jan

```
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       44; Conservative
A; Reference number: 221970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                          A; Molecule type: DNA
A; Residues: 1-170 <LYN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
A; Residues: 1-510 <MUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Accession: S56233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Gene: SGD:GAT1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
S53377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Rylangdon, T.; Savegarents, A.; Ravagnani, A.; Gielkens, M.; Caddick, M.X.; Arst Jr., H.N.
Mol. Microbiol. 17, 877-888, 1995
A; Title: Mutational analysis reveals dispensability of the N-terminal region of the Aspetite: Mutational analysis reveals dispensability of the N-terminal region of the Aspetacence number: S70167; MUID:96123430; PMID:8596437
A; Reference number: S70168
A; Retaus: nucleic acid sequence not shown
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-791, A', 793, T', 795, 'SPGTNS', 802-876 <LAN>
A; Residues: 1-876 <CAD>
A; Residues: 1-876 <CAD>
A; Cross-references: EMBL:X52491; NID:91019911; PIDN:CAA36731.1; PID:91154625
                                                                                                                                                                                                                                                                                                                                      regulatory protein areA - Emericalla nidulans
C;Species: Emericella nidulans, Aspergillus nidulans
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 16-Jun-2000
C;Accession: A57988; S10017; S70168; S72883
R;Kudla, B.; Caddick, M.X.; Langdon, T.; Martinez-Rossi, N.M.; Bennett, C.F.; Sibley, S.EMBO J. 9, 1355-1364, 1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     probable nitrogen regulatory transcription factor gafl - fission yeast (Schizosaccharomy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A,Title: The regulatory gene areA mediating nitrogen metabolite repression in Aspergillu A,Reference number: $10017; MUID:90228331; PMID:1970293
A,Accession: A57988
A,Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Description: mediates nitrogen metabolite repression C;Superfamily: nitrogen regulatory protein nit-2; GATA-type zinc finger homology C;Keywords: DNA binding; transcription regulation; zinc finger F;670-723/Domain: GATA-type zinc finger homology <GZE> F;673-697/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Cross-references: EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g1154625 A;Note: this sequence represents reinterpretation to include two exons A;Accession: S10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C.Species: Schizosaccharomyees pombe
C; Species: Schizosaccharomyees pombe
C; Date: 20-Oct-2000 #sequence_revision 08-Dec-2000 #text_change 08-Dec-2000
C; Accession: T41336; T41209, T43290
C; Accession: T41336; T412096; T43297
R; Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, January 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                   Gaps
                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: DNA
A; Residues: 158-876 <KU2>
A; Cross-references: EMBL:X52491
A; Note: this sequence represents the authors, original translation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 277; DB 1; Length 876; Pred. No. 2.9e-27; 1. Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   673 CINCFIQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 722
   Score 277; DB 2; Length 862;
Pred. No. 2.8e-27;
.; Mismatches 0; Indels
                                                                                                                                                              662 CINCFIQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKIDVIKKRNR 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                              1 CTNCFTOTTPLWRRNPDGOPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1;
                                                          1;
   98.6%;
98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98.6%;
ilarity 98.0%;
Conservative
                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: DNA
A; Residues: 1-876 <KUD>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Gene: areA
A;Introns: 147/2
                                                                   49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Function:
                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           T41336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                 δ
                                                                                                                                                                                        d
```

```
A; Molecule type: DNA
A; Residues: 129-855 <SEE>
A; Residues: 129-855 <SEE>
A; Cross-references: BMBL:AL049521; NID:g6318249; PIDN:CAB40003.1; PID:g4539591; GSPDB
A; Cross-references: BMBL:AL049521; NID:g6318249; Rim, D.U.; Jang, Y.J.; Yoo, O.J.; Yo
Gene 215, 319-328, 1998
A; Title: Molecular cloning of gaf1, a Schizosaccharomyces pombe GATA factor, which ca
A; Reference number: Z22401; MUID:98382525; PMID:9714831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          probable membrane protein YFL021w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Species: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
C;Accession: S56233
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces A;Reference number: S56186
                                                                       A;Residues: 1-170 <LYN>
A;Cross-references: EMBL:AL035076; NID:94107277; PIDN:CAA22647.1; PID:94107278; GSPDB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Residues: 1-510 <MUR>
A;Cross-references: EMBL:D50617; NID:g836685; PID:d1009858; PID:g836733; MIPS:YFL021w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 566-855 <HOE>
A;Residues: EMBL:L31601; NID:g710606; PIDN:AAC35593.1; PID:g710607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Map position: 3R
C;Superfamily: GATA-type zinc finger homology
C;Keywords: transcription factor; transcription regulation; zinc finger
F;632-685/Domain: GATA-type zinc finger homology <GZF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                A.Experimental source: strain 972h.; cosmid c417
R.Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, March 1999
A.Reference number: Z21928
A.Accession: T41208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 231; DB 2; Length 510;
Pred. No. 1.3e-21;
4; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 2; Length 855;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Map position: 6L
C;Superfamily: GATA-type zinc finger homology
C;Keywords: transmembrane protein; zinc finger
F;14-30/Domain: transmembrane #status predicted <TMI>F;307-360/Domain: GATA-type zinc finger homology <G2F>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 252; DB 2;
Pred. No. 4.4e-24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Gene: SPDB:SPCC417.01c; SPDB:SPCC1902.01; gaf1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Cross-references: SGD:S0001873; MIPS:YFL021w
A; Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82.2%;
80.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89.7%;
88.0%;
```

```
A; Residues: 1-269 CURN>
A; Cross-references: EMBL:228259; NID:g486468; PID:g486469; MIPS:YKR034w
A; Cross-references: Extain S288C
A; Experimental source: strain S288C
R; Coornaert, D.; Vissers, S.; Andre, B.; Grenson, M.
Curr. Genet. 21, 301-307, 1992
A; Title: The UGA43 negative regulatory gene of Saccharomyces cerevisiae contains both A; Reference number: $26868; MUID:92405226; PMID:1525858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted to the EMBL Data Library, December 1994
A;Description: The sequence of S. cerevisiae cosmids 9379, 9581, and lambda clone 467
A;Reference number: 850536
A;Accession: 850536
A;Accession: 850536
A;Accession: 50 OIE>
A;Residue: 1-730 OIE>
A;Residues: 1-730 OIE>
B;Rinchart, P.L.; Magasanik, B.
Mol. Cell. Biol. 11, 6216-6228, 1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Title: Sequence and expression of GLN3, a positive nitrogen regulatory gene of Sacc A; Reference number: S22280; MUID:92049353; PMID:1682800 A; Accession: S22280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Species: Saccharomyces cerevisiae
C;Date: 28-40a-1995 #sequence_revision 12-May-1995 #text_change 29-Oct-1999
C;Accession: S50543; S22280
R;Dietrich, F.S.
                                                                        A; Cross-references: EMBL:M77821; NID:g171371; PIDN:AAA34556.1; PID:g171372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A,Molecule type: DNA
A,Residues: 1-5,'L',7-206,'I',208-269 <COO>
A;Cross-references: EMBL:X60199; NID:94751; PIDN:CAA42757.1; PID:94752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A,Gene: SGD:DAL80; UGA43
A,Cross-references: SGD:S0001742; MIPS:YKR034w
A,Map position: 11R
C,Superfamily: GATA-type zinc finger homology
C,Keywords: DNA binding; nucleus; transcription factor; zinc finger
F;28-81/Domain: GATA-type zinc finger homology <GZF>
F;31-55/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 72.6%; Score 204; DB 2; Length 269; Best Local Similarity 74.0%; Pred. No. 2e-18; Matches 37; Conservative 2; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 71.5%; Score 201; DB 2; Length 730; Best Local Similarity 75.0%; Pred. No. 1.2e-17; Matches 36; Conservative 4; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C; Keywords: Zinc finger
F;303-356/Domain: GATA-type Zinc finger homology <GZF>
F;306-330/Region: zinc finger GATA motif
                                                                                                     R,Urrestarazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Status: not compared with conceptual translation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       protein - yeast (Saccharomyces cerevisiae)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Cross-references: SGD:S0000842; MIPS:YER040w
A;Map position: 5R
C;Superfamily; GATA-type zinc finger homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-473, 'G', 475-730 <MIN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N;Alternate names: protein YER040w
                                                                                                                                                                         A; Reference number: S38097
   A; Molecule type: DNA
A; Residues: 1-269 <CUN>
                                                                                                                                                                                                                                                  A; Molecule type: DNA
                                                                                                                                                                                                             A; Accession: S38106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Accession: S26868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: SGD:GLN3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLN3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 transcription factor DAL80 - yeast (Saccharomyces cerevisiae)
NyAlternate names: protein YKR034w; transcription factor UGA43; zinc finger protein UGA4
Syspeciaes: Saccharomyces cerevisiae
C;Speciaes: Saccharomyces cerevisiae
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C;Accession: 522781; S38106; S26868; S32602
R;Cunningham, T.S.; Cooper, T.G.
Mol. Cell. Biol. 11, 6205-6215, 1991
A/Title: Expression of the DAL80 gene, whose product is homologous to the GATA factors a trogen catabolite repression.
A/Reference number: S22781; MUID:92049352; PMID:1944286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Molecule type: DNA
A;Residues: 1-551 <RAW>
A;Cross-references: EMBL:249385; NID:91008294; PIDN:CAA89405.1; PID:91008295; MIPS:YJL11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Cross-references: EMBL:249385; NID:g1008294; PIDN:CAA89405.1; PID:g1008295; MIPS:YJL11
R;Boudekou, S.; Jauniaux, J.C.; Czlepluch, C.; Andre, B.
submitted to the EMBL Data Library, April 1995
A;Reference number: S57994
                                                                                                                                                                                        submitted to the EMBL Data Library, February 1995
A:Description: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and
A:Reference number: S53376
A:Accession: S53377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R;Rasmussen, S.W.
Yeast 11, 873-883, 1995
A;Title: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and CSD3
A;Title: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and CSD3
A;Reference number: S57357; MUID:96090136; PMID:7483851
A;Accession: S57358
A;Status: nucleic acid sequence not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
GZF3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J0806; protein YJL110c
C;Species: Saccharomyces cerevisiae
C;Date: 05-Nay-1995 #sequence_revision 01-sep-1995 #text_change 29-Oct-1999
C;Accession: 553377; 556888; 556891; 557994; 557358
                                                                                                                                                                                                                                                                                                                                                                                       A; Cross-references: EMBL: X85021; NID: 9728698; PIDN: CAA59384.1; PID: 9728700 R; Rasmussen, S.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Cross-references: EMBL; X86353; NID: 9899480; PIDN: CAA60126.1; PID: 9899481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14.7%; Score 210; DB 2; Length 551; ilarity 76.0%; Pred. No. 6.6e-19; Conservative 2; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ricziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C. submitted to the Protein Sequence Database, September 1995 A; Reference number: S56891
                                                                                                                                                                                                                                                                                                                                                                                                                                                        R;Rasmussen, S.W.
submitted to the Protein Sequence Database, September 1995
A;Reference number: $56876
A;Accession: $56888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Keywords: zinc finger
F; 128-181/Domain: GATA-type zinc finger homology <GZF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Cross-references: SGD:S0003646; MIPS:YJL110c
A; Map position: 10L
C; Superfamily: GATA-type zinc finger homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
A; Residues: 101-350;401-551 <RAF>
A; Cross-references: EMBL:X85021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
hes 38; Conserv
                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA
A; Residues: 1-551 <RAS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-551 <CZI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: DNA
A; Residues: 1-551 <BOU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S22781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: S56891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: S57994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Gene: SGD:GZF3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

ö

```
A; Reference number: S40382; MUID:94244465; PMID:8187633
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31; Conservative
                                   A; Accession: S40382
A; Molecule type: mRNA
A; Residues: 1-779 <ABE>
                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Residues: 1-391 < LAV>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Accession: 150702
                                                                                                                                                                                                     A; Gene: FlyBase:srp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: 150703
                                                                                                                                                                                                                                                                                                                                                                                                                                                               32;
                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                      C; Species: Xenopus laevis (African Clawed frog)
C; Date: 13.5ep-1996 #sequence_revision 13.5ep-1996 #text_change 08-Dec-2000
C; Accession: 151420
C; Accession: 151420
B; Kelley, C.; Blumberg, H.; Zon, L.I.; Evans, T.
Development 118, 817-827, 1993
A; Title: GATA-4 is a novel transcription factor expressed in endocardium of the developing A; Reference number: 151419; MUID: 94357077; PMID: 8076520
A; Accession: 151420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Abel, T.; Michelson, A.M.; Maniatis, T.
Development 119, 623-633, 1993
A;Title: A Drosophila GATA family member that binds to Adh regulatory sequences is expre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Accession: 151419
R;Kelley, C.; Blumberg, H.; Zon, L.I.; Evans, T.
Development 118, 817-827, 1993
A;Title: GATA-4 is a novel transcription factor expressed in endocardium of the developity A;Reference number: 151419; MUID:94357077; PMID:8076520
A;Accession: 151419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          transcription factor xGATA-4a - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 08-Dec-2000
C;Accession: I51419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  box A-binding factor - fruit fly (Drosophila melanogaster)
N;Alternate names: ABF: transcription factor dGATAb
C;Species: Drosophila melanogaster
C;Species: 13-Van-1995 #sequence_revision 06-Sep-1996 #text_change 16-Jul-1999
C;Accession: $40382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: mRNA
A; Residues: 1-388 < KEL>
A; Coss-references: GB:L13702; NID:g311051; PIDN:AAA63687.1; PID:g311052
C; Superfamily: transcription factor GATA-4; GATA-type zinc finger homology
C; Keywords: zinc finger
F;181-234/Domain: GATA-type zinc finger homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Residues: 17390 <KREL>
A; Residues: 17300 <KREL>
A; Cross-references: GB:L13701; NID:g311049; PIDN:AAA63686.1; PID:g311050
C; Superfamily: transcription factor GATA-4; GATA-type zinc finger homology
C; Keywords: zinc finger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        138 CINCHISTILWRRNSEGEPVCNACGLYMKLHGVPRPLAMKKESIQTRKR 287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      69.4%; Score 195; DB 2; Length 390; 64.0%; Pred. No. 3.9e-17;
live 9; Mismatches 9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 195; DB 2; Length 388; Pred. No. 3.9e-17; 9; Mismatches 9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                             1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;181-234/Domain: GATA-type zinc finger homology <GZF1>F;235-288/Domain: GATA-type zinc finger homology <GZF2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;180-233/Domain: GATA-type zinc finger homology <C2F1> F;234-287/Domain: GATA-type zinc finger homology <C2F2>
                                                                                                                                                                                                  transcription factor xGATA-4b - African clawed frog
                                                                                                                                                                                                                                                                                                                                                                                                                            A;Accession: Ib1420
A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Status: preliminary; translated from GB/EMBL/DDBJ A;Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           69.4%;
64.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 64.09
Matches 32; Conseryative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                          RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     I51419
                                                       g
δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
```

```
A;Cross-references: EMBL:X76217; NID:9441491; PIDN:CAA53807.1; PID:9441492 C;Comment: This transcriptional activator is the earliest known marker of the develop C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cross-references: EMBL:U11888; NID:9511481; PIDN:AAA57504.1; PID:9511482

G; Comment: This transcriptional activator is named for the core, GATA, of the nucleot G; Comment: This transcriptional activator is named for the core, GATA, of the nucleot G; Genetics:

A; Genetics:
C; Superfamily: transcription factor GATA-4; GATA-type zinc finger homology C; Steywords: DNA binding; nucleus; transcription factor; transcription regulation; zin F; 183-235/Domain: GATA-type zinc finger homology <GZFI>
F; 236-240/Region: zinc finger GATA motif
F; 239-263/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      transcription factor GATA-6 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 26-Aug-1999
C;Accession: ISO703
R;Laverriere, A.C.; MacNeill, C.; Mueller, C.; Poelmann, R.E.; Burch, J.B.; Evans, T.
J Biol. Chem. 269, 23177-23184, 1994
A;Title: GATA-4/5/6, a subfamily of three transcription factors transcribed in develo
A;Reference number: A54720; MUID:94365018; PMID:8083322
                                                                                                                                                                                                                                                         zin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ricaverriere, A.C.; MacNeill, C.; Mueller, C.; Poelmann, R.E.; Burch, J.B.; Evans, T. J. Biol. Chem. 259, 23177-23184, 1994
A;Title: GATA-4/5/6, a subfamily of three transcription factors transcribed in develo A;Reference number: A54720; MUID:94365018; PMID:9083222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 1-387 <LAV>
A; Cross-references: EMBL:U11889; NID:9511483; PIDN:AAA57505.1; PID:9511484
C; Comment: This transcriptional activator is named for the core, GATA, of the nucleot
                                                                                                                                                          A,Cross-references: FlyBase:FBgn0003507
C;Superfamily: box A-binding factor; GATA-type zinc finger homology
C;Superdas: DNA binding; nucleus; transcription factor; transcription regulation;
F;316-369/Domain: GATA-type zinc finger homology <GZF>
F;319-343/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 26-Aug-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   239 CINCHTINTILMERNAEGEPVCNACGLYMKLHGVPRPLAMKKESIQTRKR 288
                                                                                                                                                                                                                                                                                                                                                                                                                    67.6%; Score 190; DB 1; Length 779; 64.0%; Pred. No. 3.3e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               67.3%; Score 189; DB 2; Length 391; 62.0%; Pred. No. 2.3e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: 150702
A; Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               transcription factor GATA-5 - chicken
```

```
g
QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C;Accession: S27473
R;Voisard, C.P.; Wang, J.; Xu, P.; Leong, S.A.
Submitted to the EMBL Data Library, January 1992
A;Description: Isolation and characterization of urbSl, a gene involved in the regulation A;Reference number: S27473
A;Accession: S27473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Tamura, S.; Wang, X.H.; Maeda, M.; Futai, M.
Proc. Natl. Acad. Sci. U. S.A. 90, 10875-10880, 1993
A;Title: Gastric DN-binding proteins recognize upstream sequence motifs of parietal cell A;Reference number: A49645; MUID:94068504; PMID:8248184
                        A Gener GATA-6
C. Superfamily: transcription factor GATA-4, GATA-type zinc finger homology
C. Superfamily: transcription factor; transcription factor; transcription regulation; zinc
C. Keywords: DNA binding; nucleus; transcription factor; transcription regulation; zinc
F:178-231/Domain: GATA-type zinc finger homology GGFI>
F:231-285/Domain: GATA-type zinc finger homology GGFE>
F:235-285/Domain: Zinc finger GATA motif
F:235-259/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              transcription factor GATA - rat
N;Alternate names: DNA binding protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 08-Dec-2000
C;Accession: A49645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Species: Ustilago maydis (corn smut)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Gene: GATA-GT1
C:Superfamily: transcription factor GATA-4; GATA-type zinc finger homology
C:Keywords: zinc finger
F:185-238/Domain: GATA-type zinc finger homology <GZFI>
F:239-292/Domain: GATA-type zinc finger homology <GZF2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                     0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66.5%; Score 187; DB 2; Length 950; 64.0%; Pred. No. 9.5e-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66.9%; Score 188; DB 2; Length 391; 62.0%; Pred. No. 3.1e-16; tive 9; Mismatches 10; Indels
                                                                                                                                                                                                                                                           66.9%; Score 188; DB 2; Length 387;
62.0%; Pred. No. 3.1e-16;
Live 9; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: DNA
A; Residues: 1-950 <VOI>
A; Cross-references: EMBL:M80547; NID:q295416; PID:q170596
C; Superfamily: GATA-type zinc finger homology
C; Keywords: DNA binding; transcription regulation; zinc finger
F; 335-388/Domain: GATA-type zinc finger homology <GZF1>
F; 479-532/Domain: GATA-type zinc finger homology <GZF2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Residues: 1-391 <TM>
A;Cross-references: GB:L22760; NID:g437667; PID:g437668
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: A49645
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  URBS1 protein - smut fungus (Ustilago maydis)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 62.09
Matches 31, Conservative
                                                                                                                                                                                                                                                                                          Best Local Similarity 62.0%
Matches 31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 32; Conserv
                                                                                                                                                                                                                                                                 Query Match
      C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
S27473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                  δ
```

Search completed: January 3, 2003, 19:51:45 Job time: 25 secs

|   | • |   |  |   |   |
|---|---|---|--|---|---|
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
| • |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   | • |   |  |   |   |
|   |   |   |  | • |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   | · |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   |   |
|   |   |   |  |   | · |
|   |   |   |  |   | • |

```
January 3, 2003, 19:48:43; Search time 187.5 Seconds (without alignments) 171.929 Million cell updates/sec
                                                                                                                                                                                                                                             281
1 CINCFIQTTPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                              4569144
GenCore version 5.1.3 Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                             4569144 seqs, 644733110 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Listing first 45 summaries
                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pending_Patents_AA_Main:*
                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0% Maximum Match 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                    US-09-725-010-3
                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                Perfect score:
                                                                                                                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                                                                                  Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                                                                  Run on:
```

/cgg2\_6/ptodata/1/paa/US094\_COMB.pep: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep: / cgn2\_6/ptodata/1/paa/US080\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US083\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US084\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*
/ cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\* /cgn2\_6/ptodata/1/paa/US091\_COMB.pep: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep:\*/cgn2\_6/ptodata/1/paa/US06\_COMB.pep:\*/cgn2\_6/ptodata/1/paa/US07\_COMB.pep:\* beb: /cgn2\_6/ptodata/1/paa/US090\_COMB.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

|        |       |       |                 |    | SUMMARIES            |                  |      |
|--------|-------|-------|-----------------|----|----------------------|------------------|------|
|        |       | æ     |                 |    |                      | ,                |      |
| Result |       | Query |                 |    |                      |                  |      |
| No.    | Score | Match | Match Length DB | DB | ID                   | Description      |      |
| 1      | 281   | 100.0 | 1036            | 21 | US-09-791-537-26785  | Sequence 26785,  | 85,  |
| 7      | 281   | 100.0 | 1036.           | 21 | US-09-791-537-121036 | Sequence 121036, | 036  |
| m      | 277   | 98.6  |                 | 21 | US-09-791-537-44327  |                  | 27,  |
| 4      | 277   | 98.6  |                 | 21 | US-09-791-537-118817 | Sequence 118817, | 1817 |
| വ      | 277   | 98.6  |                 | 21 |                      | Sequence 125732, | 732  |
| 9      | 277   | 98.6  |                 | 21 | US-09-791-537-116651 | Sequence 1166    | 651, |

. ∢

g

| 7 8  | 277 | 98.6 | 865 | 21 | US-09-791-537-118818 | Sequence 118818,  |
|------|-----|------|-----|----|----------------------|-------------------|
| ) O  | 277 | 98.6 | 866 | 21 | -09-791-537-449      | Sequence 44920, A |
| 10   | 277 | 98.6 | 998 | 21 | -09-791-5            |                   |
| 11   | 277 | 98.6 | 876 | 18 | -09-487-5            |                   |
| 12   | 277 | 98.6 | 876 | 18 | -09-487-558          | 44,               |
| 13   | 277 | 98.6 | 876 | 21 | -09-791-5            | 175(              |
| 14   | 277 | 98.6 | 876 | 21 | -09-791-537          | 47207             |
| 15   | 277 | 98.6 | 876 | 22 | -09-801-368-44       |                   |
| 16   | 277 | 98.6 | 882 | 21 | -09-791-             | 81931             |
| 17   | 277 | 98.6 | 944 | 21 | -09-791-             |                   |
| 18   | 277 | 98.6 | 971 | 21 | -09-791-537-1512     |                   |
| 19   | 274 | 97.5 | 99  | 21 | -09-791-             |                   |
| 20   | 269 | 95.7 | 926 | 21 | US-09-791-537-118828 | 1186              |
| 21   | 256 | 91.1 | 688 | 24 | -10-032-             |                   |
| 22   | 256 | 91.1 | 719 | 16 | ç                    | 17559             |
| 23   | 256 | 91.1 | 719 | 25 | -10-179              |                   |
| . 24 | 256 | 91.1 | 719 | 27 | US-60-096-409-17559  | 17559             |
| 25   | 252 | 89.7 | 290 | 21 | -09-791-537-11882    |                   |
| 26   | 252 | 89.7 | 727 | 21 | US-09-791-537-148143 |                   |
| 27   | 231 | 82.2 | 156 | 21 | US-09-791-537-89028  |                   |
| 28   | 231 | 82.2 | 510 | 21 | US-09-791-537-118210 | Sequence 118210,  |
| 29   | 212 | 75.4 | 205 | 16 | US-09-248-796-17564  | 17564,            |
| 30   | 212 | 75.4 | 202 | 27 | 960-0                | 1756              |
| 31   | 212 | 75.4 | 720 | 25 | -10 - 179            | 9251              |
| 32   | 210 | 74.7 | 551 | 19 | 9-538                | 448,              |
| 33   | 206 | 73.3 | 694 | 25 | -179-131-5           | 5254, A           |
| 34   | 204 | 72.6 | 135 | 21 | -09-791              | 7904              |
| 35   | 204 | 72.6 | 192 | 18 | -0                   | 28191,            |
| 36   | 204 | 72.6 | 269 | 21 | -09-791-537-7384     | equence 7384      |
| 37   | 204 | 72.6 | 569 | 21 | 3-791                | ednence 80768     |
| 38   | 204 | 72.6 | 298 | 18 | -09-487              | 230,              |
| 39   | 204 | 72.6 | 298 | 18 | -09-487-558B-        | equence 230,      |
| 40   | 204 | 72.6 | 298 | 22 | -368-23              |                   |
| 41   | 201 | 71.5 | 730 | 18 | 558-12               |                   |
| 42   | 201 | 71.5 | 730 | 18 | ņ                    | Sequence 126, App |
| 43   | 201 | 71.5 | 730 | 22 | -801-368-1           |                   |
| 44   | 195 | 69.4 | 383 | 21 | US-09-791-537-46380  |                   |
| 45   | 195 | 69.4 | 388 | 21 | US-09-791-537-29921  | Sequence 29921, A |
|      |     |      |     |    |                      |                   |

#### ALIGNMENTS

US-09-791-537-26785; Sequence 26785, Application US/09791537; GENERAL INFORMATION:

```
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY-ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE PERENENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 26785
LENGTH: 1036
                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 281; DB 21; Length 1036; 100.0%; Pred. No. 3.8e-27; ive 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                    ; TYPE: PRT
; ORGANISM: Neurospora crassa
US-09-791-537-26785
                                                                                                                                                                                                                                                                                                                                                                                                        50; Conservative
APPLICANT: Bionomix, Inc. APPLICANT: Debe, Derek
                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 50; Conserv
                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ōλ
```

```
US-09-791-537-125732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-791-537-125732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-791-537-116651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-791-537-116651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 125732
LENGTH: 860
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQ ID NO 116651
LENGTH: 862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         49;
      LENGTH: 725
                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
                             TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
                                                                                                                                                                                              δ
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QΥ
                                                                                                APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Debe, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION UNMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTHARE: PATENTIN VERSION 3.0
SEQ ID NO 121036
LENGTH: 1036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 44327, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Blonomix, Inc.
APPLICANT: Blonomix, Inc.
APPLICANT: Blonomix, Inc.
APPLICANT: Dabe, Derek
APPLICANT: Dabe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: HERED DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT PILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PATCHIN VERSION 3.0
SEQ ID NO 118817
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0;
                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1036;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 277; DB 21; Length 66;
Pred. No. 6.3e-28;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 100.0%; Score 281; DB 21; Best Local Similarity 100.0%; Pred. No. 3.8e-27; Matches 50; Conservative 0; Mismatches 0;
                                        Sequence 121036, Application US/09791537 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-791-537-118817; Sequence 118817, Application US/09791537; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98.68;
                                                                                                                                                                                                                                                                                                                                                      ; ORGANISM: Neurospora crassa
US-09-791-537-121036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                   APPLICANT: Bionomix, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Bionomix, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; ORGANISM: pdb 4GATA
US-09-791-537-44327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
RESULT 2
US-09-791-537-121036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
US-09-791-537-44327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQ ID NO 44327
LENGTH: 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REPERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: BLONGLY.
APPLICANT: BLONGLY.
APPLICANT: BLONGLY.
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: MFTHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PATENTIN version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ő
                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                       Length 725;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 860;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98.6%; Score 277; DB 21; Length 862; 98.0%; Pred. No. 1e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         660 CINCFIQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 709
                                                                                                                                                                                                   525 CINCFTQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                     Score 277; DB 21; Length 7;
Pred. No. 8.6e-27;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 277; DB 21; Length 8:
Pred. No. 1e-26;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
                                                                                                                                                                                                                                                                                                                                           Sequence 125732, Application US/09791537; GENERAL INFORMATION: APPLICANT: Bionomix, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 116651, Application US/09791537; GENERAL INFORMATION:
; ORGANISM: Penicillium chrysogenum
US-09-791-537-118817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Penicillium chrysogenum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE: PRT ORGANISM: Penicillium roqueforti
                                                                       Query Match 98.6%;
Best Local Similarity 98.0%;
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98.6%;
98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 98.6'
Best Local Similarity 98.0'
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                             Debe, Derek
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
```

```
APPLICANT: BLOOMLY, Inc.
APPLICANT: BLOOMLY, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Bloomix, Inc.
APPLICANT: Bloomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
SOFTWARE: DATE: 2001-02-22
SOFTWARE: PATENTING PATE: 2001-03-03
SECTION 08 1933
LENGTH: 866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.6%; Score 277; DB 21; Length 866; 98.0%; Pred. No. 1e-26; Live 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 98.6%; Score 277; DB 21; Length 866; Best Local Similarity 98.0%; Pred. No. 1e-26; Matches 49; Conservative 1; Mismatches 0; Indels
          664 CINCFTQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                           Sequence 44920, Application US/09791537 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 81933, Application US/09791537 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Sequence 44, Application US/09487558; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Aspergillus parasiticus US-09-791-537-44920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; ORGANISM: Aspergillus oryzae US-09-791-537-81933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 98.0 Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Madden, Kevin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Maxon, Mary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Milne, Todd
                                                                                                    US-09-791-537-44920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
US-09-791-537-81933
                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 44920
LENGTH: 866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-487-558-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                              APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REPERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 2, Application US/08750458
GENERAL INFORMATION:
APPLICANT: Christensen, Tove
APPLICANT: Hynes, Michael J.
TITLE OF INVENTION: A Fungus Wherein The AreA Gene Has Been Modified And An AreA
NUMBER OF SEQUENCES: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 277; DB 21; Length 865;
Pred. No. 1e-26;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98.6%; Score 277; DB 11; Length 866; 98.0%; Pred. No. 1e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOFTWARE: Petentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
RPPLICATION NUMBER: US/08/750,458
FILLING DATE: 3-December-1996
CLASSIFICATION: 435
ATTORNEY AGENT INFORMATION:
NAME: ROZEK, CATOL E.
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4129.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0228
TELEFAX: 212 867 0228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE: Novo Nordisk of North America, Inc. STREET: 405 Lexington Avenue, Suite 6400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Mismatches
                    Sequence 118818, Application US/09791537 GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                       ; ORGANISM: Penicillium urticae US-09-791-537-118818
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 98.6%;
Best Local Similarity 98.0%;
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      866 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 98.0%
Matches 49;, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; MOLECULE TYPE: protein US-08-750-458-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           I: New York
PRY: U.S.A.
10174-6401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New York
US-09-791-537-118818
                                                                                                                                                                                                                                                                                                              SEQ ID NO 118818
LENGTH: 865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-750-458-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

Norman, Thea

APPLICANT:

```
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
              ; Sequence 17502, Application US/09791537 ; GENERAL INFORMATION: APPLICANT: Bionomix, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; Sequence 47207, Application US/09791537; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-801-368-44; Sequence 44, Application US/09801368; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                      98.6%;
98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.6%;
ilarity 98.0%;
Conservative
                                                                                                                                                                                                                                                                                                      TYPE: PRT ORGANISM: Emericella nidulans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Emericella nidulans
                                                                                                                                                                                                                                                                                                                                                                                                                              49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Norman, Thea
Royer, John
Salama, Sofie
Sherman, Amir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Maxon, Mary
Milne, Todd
                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 49; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 49; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
US-09-791-537-47207
  JS-09-791-537-17502
                                                                                                                                                                                                                                                                                                                                                  US-09-791-537-17502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-791-537-47207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NO 47207
LENGTH: 876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Madden, Kevin
APPLICANT: Madden, Kevin
APPLICANT: Mann, Mary
APPLICANT: Milne, Todd
APPLICANT: Milne, Todd
APPLICANT: Milne, Todd
APPLICANT: Salma, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Summers, Eric
CURRENT: APPLICATION: Webbers: US/09/487,558B
CURRENT APPLICATION NUMBER: US 60/487,558
PRIOR APPLICATION NUMBER: US 60/487,558
PRIOR APPLICATION NUMBER: US 60/487,558
PRIOR APPLICATION NUMBER: US 60/487,558
SPRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 446
SSOFTWARE: Patentin version 3.0
SEQ ID NO 44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                       APPLICANT: SITUE OF INVENTION: Methods for Improving Secondary Metabolite Production TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production TITLE OF INVENTION: in Fungi
FILLE REPERENCE: 109272.147
CURRENT PAPLICATION NUMBER: US/09/487,558
CURRENT FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 1999-10-20
NUMBER: OF SEQ ID NOS: 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             673 CTNCFTQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 277; DB 18;
Pred. No. 1.1e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 277; DB 18;
Pred. No. 1.1e-26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 44, Application US/09487558B GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                         , ORGANISM: Aspergillus nidulans US-09-487-558-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM: Aspergillus nidulans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98.6%;
98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holtzman, Doug
Madden, Kevin
                         Sherman, Amir
    Salama, Sofie
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Busby, Robert APPLICANT: Cali, Brian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hecht, Peter
                                            Silva, Jeff
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 12
US-09-487-558B-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-487-558B-44
                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 44
LENGTH: 876
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
```

```
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT PILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 17502
LENGTH: 876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 876;
                                                                                                                                                                                                                                                                                                                                 Length 876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CINCFIQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 277; DB 21;
Pred. No. 1.1e-26;
1; Mismatches 0;
                                                                                                                                                                                                                                                                                                                               Score 277; DB 21;
Pred. No. 1.1e-26;
```

```
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITE OF INVENTION Methods for Improving Secondary Metabolite Production in Fungi
FILE FREERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2000-03-07
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PAGENTAL OF SEQ ID NOS: 440
LUMBER OF SEQ ID NOS: 440
SOFTWARE: PAGENTAL OF SEQ ID NOS SEQ ID NOS 440
LUMBER OF SEQ ID NOS 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
98.6%; Score 277; DB 22; Length 876;
Best Local Similarity 98.0%; Pred. No. 1.1e-26;
Matches 49; Conservative 1; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: January 3, 2003, 19:59:15 Job time : 188.5 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT (CRANISM: Aspergillus nidulans US-09-801-368-44)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

| · |  |  |
|---|--|--|

À

```
January 3, 2003, 19:50:43; Search time 19.5 Seconds (without alignments) 181.016 Million cell updates/sec
                                                                                                                                                                                                                                                        US-09-725-010-3
281
1 CINCFIQITPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                258925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pending_Patents_AA_New:*

1: /cgn2_6/ptodata/2/paa/USO6_NEW_COMB.pep:*

2: /cgn2_6/ptodata/2/paa/USO6_NEW_COMB.pep:*

4: /cgn2_6/ptodata/2/paa/USO7_NEW_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/USO8_NEW_COMB.pep:*

5: /cgn2_6/ptodata/2/paa/USO8_NEW_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/USO8_NEW_COMB.pep:*

7: /cgn2_6/ptodata/2/paa/USO8_NEW_COMB.pep:*
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                             258925 seqs, 70596210 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                           Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                             Searched:
                                                                                                                                                         Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Description         | Sequence 3, Appli | Sequence 2, Appli | Sequence 87510, A | Sequence 87510, A | Sequence 92485, A | Sequence 92485, A | Sequence 92294, A | Sequence 92294, A | Sequence 87508, A | Sequence 87508, A | Sequence 87509, A   | Sequence 87509, A    | Sequence 4, Appli | 98        | 32               |      | 17    | 61       | 61656, | 61641, | 61641, | 6171       | 61710,     | 61679,        | 61679,               |                     |
|---------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|---------------------|----------------------|-------------------|-----------|------------------|------|-------|----------|--------|--------|--------|------------|------------|---------------|----------------------|---------------------|
| ID                  | -09-725-          | -725-             | -724-676-         | -09-724-          | -09-724-676-      | -724-67           | -09-724-676-      | -724-67           | 09-724-           | -724              | US-09-724-676-87509 | US-09-724-676A-87509 | 627-              | 386-264-9 | US-10-278-536-32 | 171- | 302-3 | -09-724- | -724-  | -724-  | -724   | 4-676-6171 | -724       | -724-676-6167 | US-09-724-676A-61679 | US-09-724-676-61663 |
| DB                  | r.                | Ŋ                 | Ŋ                 | 'n                | 2                 | 'n                | Ŋ                 | Ŋ                 | Ŋ                 | 'n                | Ŋ                   | S                    | 9                 | 9         | 9                | 9    | -     | Ŋ        | S      | Ŋ      | 2      | S          | 2          | Ŋ             | Ŋ                    | S                   |
| Length              | 50                | 20                | 347               | 347               | 357               | 357               | 455               | 455               | 203               | 203               | 471                 | 471                  | 291               | 291       | 139              | 139  | 593   | 267      | 267    | 270    | 270    | 301        | 301        | 304           | 304                  | 372                 |
| %<br>Query<br>Match |                   |                   |                   |                   |                   | 64.8              |                   | 63.3              | 45.9              | 45.9              | 45.9                | 5                    | 5                 | 5.        | ÷                | ä    | 3.    | ά.       | ς.     | ζ.     | ď      | 7          | ζ.         | 3             | 22.2                 | 2                   |
| Score               | 281               | 256               | 188               | 188               | 182               | 182               | 178               | 178               | 129               | 129               | 129                 | 129                  | 100               | 100       | δ                | 89.5 | 99    | $\sim$   | 62.5   | $\sim$ | $\sim$ | $\sim$     | $^{\circ}$ | $\sim$        | 62.5                 | 62.5                |
| Result<br>No.       | П                 | 7                 | ю                 | 4                 | S                 | 9                 | 7                 | 80                | 6                 | 10                | 11                  | 12                   | 13                | 14        | 15               | 16   | 17    | 18       | 19     | 20     | 21     | 22         | 23         | 24            | 25                   | 26                  |

| Sequence 61663, A Sequence 61655, A Sequence 61655, A Sequence 61640, A Sequence 61640, A Sequence 61709, A Sequence 61578, A Sequence 61678, A Sequence 61678, A Sequence 61639, A Sequence 61677, A Sequence 61677, A Sequence 61677, A Sequence 61662, A Sequence 61662, A  |            | ON OF HYPHAL-SPECIFIC  | Length 50; Indels 0; Gaps 0; IKKRNR 50        IKKRNR 50                                      | ION OF HYPHAL-SPECIFIC   |
|--|------------|--|--|--|
| US-09-724-676A-61663<br>US-09-724-676-61655<br>US-09-724-676-61655<br>US-09-724-676-61440<br>US-09-724-676-61709<br>US-09-724-676-61709<br>US-09-724-676-61709<br>US-09-724-676-61678<br>US-09-724-676-61678<br>US-09-724-676A-61654<br>US-09-724-676A-61654<br>US-09-724-676A-61639<br>US-09-724-676A-61639<br>US-09-724-676-61708<br>US-09-724-676-61708<br>US-09-724-676-61677<br>US-09-724-676-61677<br>US-09-724-676-61677<br>US-09-724-676-61677<br>US-09-724-676-61677<br>US-09-724-676-61677 | ALIGNMENTS | 5/09725010  JLA  DDS FOR ALTERING THE EXPRESSION  SS  SR: US/09/725,010  10-11-29  11-29  2.1  | ; Score 281; DB 5;<br>; Pred. No. 1.3e-27;<br>0; Mismatches 0;<br>NACGLFLKLHGVVRPLSLKTDV<br> | 9725010<br>FOR ALTERING THE EXPRESSION OF US/09/725,010 11-29 60/167,672   |
| 222<br>222<br>222<br>222<br>222<br>222<br>222<br>222<br>222<br>22  |            | -3 Application US/09 Application US/09 SUNDSTROM, PAULA NVENTION: METHODS NVENTION: GENES ENCE: 23878.0005 ELICATION NUMBER: 60 ICATION NUMBER: 60 | 100.0% Similarity 100.0% 0; Conservative CETOTTEMERNPEOPLC!                                  | Application US/09 RMATION: SUNDSTROM, PAULA VENTION: METHODS NCE: 23878.0005 LICATION NUMBER: LIGATION NUMBER: CATION NUMBER: 2000-11-2 EQ ID NOS: 11 'atentin Ver: 2.1  |
| 222<br>223<br>333<br>333<br>333<br>333<br>333<br>333<br>333<br>333   |            | 1 725-010-3 ence 3, A ence 3, A EAL INFOR LICANT: 8 LE OF INV LE OF SE LE OF S | Ouery Match Best Local Simil Matches 50; C 1 CTNCFTOT  | 2<br>725-ence<br>ence<br>ENAL<br>LE C<br>LE C<br>E RENT<br>RENT<br>OOR P<br>OOR P<br>OOR P<br>OOR P<br>OOR P<br>OOR P  |
|  |            | RESULT US-09- US-09- Seque APP APP APP APP APP APP APP APP APP AP  | Ou<br>Be<br>Ma<br>Ma<br>Oy   | RESULT Sequence Seque |

```
LENGTH: 357
                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Пр
                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0;
                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                           66.9%; Score 188; DB 5; Length 347; 62.0%; Pred. No. 2.1e-15; Live 9; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 188; DB 5; Length 347;
Pred. No. 2.1e-15;
9; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
   Length 50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                         APPLICANT: Compugen LTD
TILE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTRAME: Patentin version 3.2
SOFTRAME: Patentin version 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 87510, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 87510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 92485, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SQFTWARE: PatentIn version 3.2
SEQ ID NO 92485
 Score 256; DB 5;
Pred. No. 1.5e-24;
                                1; Mismatches
                                                                                                                                                                                           Sequence 87510, Application US/09724676 GENERAL INFORMATION:
 91.1%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66.9%;
62.0%;
                                    46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Homo sapiens
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-724-676A-87510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-724-676A-87510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
US-09-724-676-92485
                                                                                                                                                                             US-09-724-676-87510
                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-724-676-87510
                                                                                                                                                                                                                                                                                                                                                                       LENGTH: 347
                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGIH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 4
                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .
0
                                                                               Score 182; DB 5; Length 357;
Pred. No. 1.2e-14;
6; Mismatches 12; Indels
                                                                                                                                                                                       258 CINCQTTTTTLWRRNASGDPVCNACGLYYKLHQVNRPLTMRKDGIQTRNR 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64.8%; Score 182; DB 5; Length 357; 64.0%; Pred. No. 1.2e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 455;
                                                                                                                                                                    1 CINCFIQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CINCFIQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 92485
LENGTH: 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURENT APPLICATION WIMBER: US/09/724,676
CURENT APPLICATION WIMBER: 28/00-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PATENTIN version 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 178; DB 5;
Pred. No. 4.5e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8; Mismatches
                                                                                                                                                                                                                                                                                                           ; Sequence 92485, Application US/09724676A; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 92294, Application US/09724676A; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Sequence 92294, Application US/09724676; GENERAL INFORMATION:
                                                                                  64.8%;
64.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63.3%;
60.0%;
                                                                             Query Match
Best Local Similarity 64.0°
Matches 32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 64.03
Matches 32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
ORGANISM: Homo sapiens
                       sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                            US-09-724-676A-92485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-724-676A-92485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
US-09-724-676A-92294
; TYPE: PRT
; ORGANISM: HOMO S
US-09-724-676-92485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-724-676-92294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-724-676-92294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 92294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 7
```

```
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                   ÷
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ĥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                          Query Match 45.9%; Score 129; DB 5; Length 203; Best Local Similarity 50.0%; Pred. No. 2.1e-08; Matches 25; Conservative 7; Mismatches 12; Indels
                                                                                                                                                                                            Length 455;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         45.9%; Score 129; DB 5; Length 203; 50.0%; Pred. No. 2.1e-08; tive 7; Mismatches 12; Indels
                                                                                                                                                                                                                                                                     1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                               12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFWRAKE: PatentIn version 3.2
SOFWRAKE: PatentIn version 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 87508, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: Patentin version 3.2
                                                                                                                                                                                          Score 178; DB 5;
Pred. No. 4.5e-14;
8; Mismatches 12;
CURRENT APPLICATION NUMBER: US/09/724,676A CURRENT FILING DATE: 2000-11-28 NUMBER OF SEQ ID NOS: 97222 SORTWARE: Patentin version 3.2 SEQ ID NO 92294
                                                                                                                                                                                                                                                                                                                                                             RESULT 9
US-09-724-676-87508
; Sequence 87508, Application US/09724676
; GENERAL INFORMATION:
                                                                                                                                                                                          63.3%;
60.0%;
                                                                                                                                                                                                                               30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25; Conservative
                                                                                                                               ; ORGANISM: Homo sapiens
US-09-724-676A-92294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; ORGANISM: Homo sapiens
US-09-724-676A-87508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Homo sapiens
                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3est_Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
US-09-724-676A-87508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-724-676-87508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQ ID NO 87508
LENGTH: 203
                                                                                              LENGTH: 455
                                                                                                                   TYPE: PRT
                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH:
                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

US-09-724-676-87509

RESULT 11

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Broun, Pierre
APPLICANT: Samaha, Raymond
APPLICANT: Zhang, James
ITILE OF INVENTION: POLYNUCLEOTIDES FOR FLOWER TRAIT ALTERATION
                                                                                                                                                                                                                                                                                                                                                        45.9%; Score 129; DB 5; Length 471; 50.0%; Pred. No. 4.6e-08; Live 7; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | || : ||||||: | ||||||| || 409 CVNCGSIQTPLWRRDGTGHYLCNACGLYSKMNGLSRPL-----IKPQKR 452
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 45.9%; Score 129; DB 5; Length 471; Best Local Similarity 50.0%; Pred. No. 4.6e-08; Matches 25; Conservative 7; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
Sequence 87509, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION UNDBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PATENTIN VERSION 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE OF INVENTION: Variants of alternative splicing FILE REFERRNCE: 129181.4 Compugen CURRENT APPLICATION WIMBER: US/09/724,676A CURRENT FILING DATE: 2000-11-28 NUMBER OF SEQ ID NOS: 97222 SOFTWARE: PALEITIN VERSION 3.2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILE REFERENCE: MBI-012
CURRENT APPLICATION NUMBER: US/10/290,627
CURRENT FILING DATE: 2002-11-07
PRIOR APPLICATION NUMBER: US/09/533,648
PRIOR FILING DATE: 2000-03-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
US-09-724-676A-87509
; Sequence 87509, Application US/09724676A
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Sequence 4, Application US/10290627; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Riechmann, Jose-Luis
APPLICANT: Heard, Jacqueline
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Heard, Jacqueline
Jiang, Cai-Zhong
Pineda, Omaira
Rueber, Lynne
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Keddie, James
Ratcliffe, Oliver
Pilgrim, Marsha
                                                                                                                                                                                                                                                                                                                                                                                                        25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yu, Guo-Liang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TYPE: PRT; ORGANISM: Homo sapiens
US-09-724-676A-87509
                                                                                                                                                                                                                                                                                    ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Adam, Luc
                                                                                                                                                                                                                                                                                                                                                                             Best_Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                            US-09-724-676-87509
                                                                                                                                                                                                                 SEQ ID NO 87509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQ ID NO 87509
LENGTH: 471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
US-10-290-627-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                            Query Match
```

1;

us-09-725-010-3.rapn

Yu, Guo-Liang

```
Dp
                                                                                                                                                                                                                                                                  4; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Broun, Dierre
APPLICANT: Reuber, Lynne
APPLICANT: Pilgrim, Marsha
APPLICANT: Samaha, Raymond
TITLE OF INVENTION: POLYNUCLEOTIDES FOR SEED TRAIT ALTERATION
FILE REPREBRUE: MBI-008
FILE REPREBRUE: MBI-008
CURRENT APPLICATION NUMBER: US/10/286,264
CURRENT APPLICATION NUMBER: 60/125,814
PRIOR PILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 165
SOSTWARE: Patentin Ver. 2.1
                                                                                                                                                                                                                      35.6%; Score 100; DB 6; Length 291; 45.7%; Pred. No. 0.0001; tive 7; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 100; DB 6; Length 291;
Pred. No. 0.0001;
7; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                      1 CTNCFTQTTPLWRRNPDG-QPLCNACGLFLKLHGVV---RPLSLKT 42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CTNCFTQTTPLWRRNPDG-QPLCNACGLFLKLHGVV---RPLSLKT 42
                                                                                                                                                FEATURE:

CTHER INFORMATION: Mendel Sequences G1504 & G1504PEP

US-10-290-627-4
PRIOR APPLICATION NUMBER: 60/125,814 PRIOR FILING DATE: 1999-03-23 NUMBER OF SEQ ID NOS: 77 SOFTWARE: PATENTIN VET: 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 15
US-10-278-536-32
; Sequence 32, Application US/10278536
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 98, Application US/10286264 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Riechmann, Jose-Luis
Ratcliffe, Oliver
                                                                                                             TYPE: PRT
ORGANISM: Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT ORGANISM: Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zhang, James
Jiang, Cai-Zhong
Pineda, Omaira
Heard, Jacqueline
Yu, Guo-Liang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              35.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 45.73
Matches 21; Conservative
                                                                                                                                                                                                                                                              21; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Adam, Luc
Broun, Pierre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; OTHER INFORMATION: G1504
US-10-286-264-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Keddie, James
                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                               US-10-286-264-98
                                                                         SEQ ID NO 4
LENGTH: 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID NO 98
LENGTH: 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURE:
                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

APPLICANT: Samaha, Raymond
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Pineda, Omaira
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose-Luis

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                            1;
                                                                                                                                                                                                                                                                                                                                                                                                  Length 139;
                                                                                                                                                                                                                                                                                                                                                                                                                                            8; Indels
                                                                       APPLICANT: Broun, Pierre
TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-011
CURRENT APPLICATION NUMBER: US/10/278,536
CURRENT APPLICATION NUMBER: 06/125,814
PRIOR PILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 238
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3: SEQ ID NOS: 238
LENGTH: 139
                                                                                                                                                                                                                                                                                                                                                                                                    Score 89.5; DB 6;
Pred. No. 0.00098;
3; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: January 3, 2003, 20:00:14
Job time: 19.5 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 CTNCFTQTTPLWRRNPDG-QPLCNACGL 27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | :| | | | | | | | | 38 CADCGTSKTPLWRGGPVGPKSLCNACGI 65
                                                                                                                                                                                                                                                                                                          ORGANISM: Arabidopsis thaliana
Keddie, James
Ratcliffe, Oliver
Pilgrim, Marsha
Adam, Luc
                                                                                                                                                                                                                                                                                                                                                                                                        31.9%;
57.1%;
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 57.18;
Matches 16; Conservative.
                                                                                                                                                                                                                                                                                                                              ; FEATURE:
; OTHER INFORMATION: G1508
US-10-278-536-32
                                       APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                    TYPE: PRT
   APPLICANT:
APPLICANT:
```

ä

```
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

January 3, 2003, 19:51:53 ; Search time 21.5 Seconds (without alignments) 44.073 Million cell updates/sec Run on:

US-09-725-010-3 281 1 CINCFIQITPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR 50 Perfect score: Sequence:

BLOSUM62 Scoring table:

Gapop 10.0 , Gapext 0.5

117078 segs, 18951520 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Listing first 45 summaries Post-processing: Minimum Match 0% Maximum Match 100%

Database:

Published\_Applications\_AA:\*

1: /cgnl2\_fytodata/1/pubpaa/USOB\_NEW\_PUB.pep:\*
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
3: /cgn2\_6/ptodata/1/pubpaa/USO6\_NEW\_PUB.pep:\*
4: /cgn2\_6/ptodata/1/pubpaa/USO6\_NEW\_PUB.pep:\*
5: /cgn2\_6/ptodata/1/pubpaa/USO7\_NEW\_PUB.pep:\*
6: /cgn2\_6/ptodata/1/pubpaa/USO7\_NEW\_PUB.pep:\*
7: /cgn2\_6/ptodata/1/pubpaa/USO7\_PUBCOMB.pep:\*
8: /cgn2\_6/ptodata/1/pubpaa/USO8\_PUBCOMB.pep:\*
9: /cgn2\_6/ptodata/1/pubpaa/USO8\_PUBCOMB.pep:\*
10: /cgn2\_6/ptodata/1/pubpaa/USO8\_PUBCOMB.pep:\*
11: /cgn2\_6/ptodata/1/pubpaa/USO9\_PUBCOMB.pep:\*
11: /cgn2\_6/ptodata/1/pubpaa/USO9\_PUBCOMB.pep:\*
13: /cgn2\_6/ptodata/1/pubpaa/USO9\_PUBCOMB.pep:\*
13: /cgn2\_6/ptodata/1/pubpaa/USO9\_PUBCOMB.pep:\*
13: /cgn2\_6/ptodata/1/pubpaa/USO0\_PUBCOMB.pep:\*
13: /cgn2\_6/ptodata/1/pubpaa/USO0\_PUBCOMB.pep:\*
14: /cgn2\_6/ptodata/1/pubpaa/USO0\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

| Description                   | Sequence 44, Appl | _`                | Sequence 126, App | Sequence 11, Appl | Sequence 360, App | Sequence 358, App | Sequence 1573, Ap  | Sequence 68, Appl | Sequence 124, App | Sequence 123, App | Sequence 46, Appl | Sequence 6, Appli | Sequence 2, Appli | Sequence 36, Appl | Sequence 5832, Ap  | Sequence 181, App | Sequence 11, Appl | Sequence 36, Appl | Sequence 8, Appli |
|-------------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|-------------------|-------------------|-------------------|
| Ω                             | US-09-801-368-44  | US-09-801-368-230 | US-09-801-368-126 | US-09-749-728B-11 | US-09-801-368-360 | US-09-801-368-358 | US-09-925-300-1573 | US-10-029-180-68  | US-10-029-180-124 | US-10-029-180-123 | US-09-801-368-46  | US-09-948-018-6   | US-09-948-018-2   | US-09-948-018-36  | US-09-738-626-5832 | US-09-764-846-181 | US-09-799-946-11  | US-09-967-552A-36 | US-10-033-026-8   |
| DB                            | 10                | 10                | 10                | 10                | 10                | 10                | 10                 | σ                 | 6                 | σ                 | 10                | 10                | 10                | 10                | 6                  | 10                | 10                | 10                | 12                |
| %<br>Query<br>Match Length DB | 876               | 298               | 730               | 442               | 532               | 549               | 128                | 359               | 440               | 451               | 588               | 136               | 355               | 380               | 1194               | 74                | 108               | 387               | 2237              |
| %<br>Query<br>Match           | 98.6              | 72.6              | 71.5              | 65.1              | 63.7              | 63.3              | 56.2               | 33.1              | 33.1              | 33.1              | 26.7              | 20.1              | 20.1              | 20.1              | 20.1               | 19.2              | 18.9              | 18.9              | 18.7              |
| Score                         | 277               | 204               | 201               | 183               | 179               | 178               | 158                | 93                | 93                | 66                | 75                | 56.5              | 56.5              | 56.5              | 56.5               | 54                | 53                | 53                | 52.5              |
| Result<br>No.                 | П                 | 7                 | e                 | 4                 | ស                 | 9                 | 7                  | æ                 | 6                 | 10                | 11                | 12                | 13                | 14                | 15                 | 16                | 17                | 18                | 19                |

| 1378<br>2, P<br>4, P                                      | Sequence 26, Appl<br>Sequence 24, Appl<br>Sequence 128, App | 52                                  | Sequence 2, Appli<br>Sequence 101, App | Sequence 1424, Ap<br>Sequence 1233, Ap   | Sequence 948, App<br>Sequence 103, App | Sequence 1118, Ap<br>Sequence 12, Appl  |  | Sequence 1, Appli<br>Sequence 4, Appli | Sequence 1352, Ap<br>Sequence 6, Appli | Sequence 18, Appl<br>Sequence 5, Appli | Sequence 18, Appl |
|---|---|-------------------------------------|--|--|--|---|--|--|--|--|-------------------|
| US-09-815-242-13782<br>US-09-824-734-2<br>US-09-900-237-4 | US-09-808-387-26<br>US-09-808-387-24<br>US-09-729-674-128   | US-09-808-387-6<br>US-09-808-387-22 | US-09-808-387-2<br>US-09-908-711-101   | US-09-764-864-1424<br>US-09-764-877-1233 | US-09-764-864-948<br>US-10-108-605-103 | US-09-867-550-1118<br>US-09-816-669A-12 | 0 US-09-880-705-2<br>2 US-10-005-947-2 | US-09-772-316-1<br>US-09-938-275-4     | US-09-925-300-1352<br>US-09-759-010-6  | US-09-991-496-18<br>US-09-759-010-5    | US-09-874-923-18  |
| 10<br>10  | 100   | 1000                                | 100                                    | 10                                       | 10                                     | 10                                      | 12                                     | 10                                     | 100                                    | 10                                     | 10                |
| 292<br>1146<br>1148                                       | 315<br>373<br>380   | 386<br>555                          | 574                                    | 188                                      | 338                                    | 87                                      | 314                                    | 3084                                   | 415<br>724                             | 732                                    | 732               |
| 18.1<br>18.1<br>18.1                                      | 18.0<br>18.0  | 118.0                               | 18.0                                   | 17.6                                     | 17.6                                   | 17.4                                    | 17.3<br>17.3                           | 17.3                                   | 17.1                                   | 17.1                                   | 17.1              |
| 51<br>51  | 50.5<br>50.5  | 50.5                                | 500.4                                  | 49.5                                     | 49.5                                   | 4<br>9<br>4<br>9                        | 48.5                                   | 48.5                                   | 4 4<br>8 8                             | 4<br>4<br>8<br>8                       | 48                |
| 20<br>21<br>22  | 2 2 2<br>2 4 5  | 226                                 | 0 0 0<br>1 0 E                         | 313                                      | 333                                    | 35<br>35                                | 37<br>38                               | 39<br>40                               | 41<br>42                               | 44                                     | 45                |

### ALIGNMENTS

```
APPLICANT: Madden, Keving
APPLICANT: Madden, Keving
APPLICANT: Maxon, Mary
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250Alman, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sharman, Amir
APPLICANT: Sharman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Silva, Jeff
APPLICANT: Sharman, Mario
APPLICANT: Silva, Jeff
APPLICANT: Somers, Eric
APPLICANT: Somers, Eric
APPLICANT: Somers, Eric
APPLICANT: Somers, Eric
APPLICANT: Mathod of Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR PLING DATE: 1999-10-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                · o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 98.6%; Score 277; DB 10; Best Local Similarity 98.0%; Pred. No. 7.9e-29; Matches 49; Conservative 1; Mismatches 0;
                                                  ; Sequence 44, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Aspergillus nidulans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NUMBER OF SEQ ID NOS: 440 SOFTWARE: PatentIn version 3.0
RESULT 1
US-09-801-368-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-801-368-44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

g

```
71.5%;
75.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ogawa, Satoshi
Sakurada, Kazuhiro
Gojo, Satoshi
                                                                                      Query Match
Best Local Similarity 75.09
                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Umezawa, Akihiro APPLICANT: Hata, Jun-Ichi
                                                                                                                                                                                                                                                                                                                                                                                                                              Fukuda, Keiichi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; ORGANISM: Homo sapiens
US-09-749-728B-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Royer, John
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
US-09-801-368-360
                                                                                                                                                                                                                                                                                                      US-09-749-728B-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID NO 11
LENGTH: 442
LENGTH: 730
                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                     οy
                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                       APPLICANT: Holtzman, Doug
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Madden, Kevin
APPLICANT: Madden, Kevin
APPLICANT: Mine, Todd
APPLICANT: Mine, Todd
APPLICANT: No. US20020128250Alman, Thea
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Summers, Eitc
TITLE OF INVERTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US 09/487,558
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           72.6%; Score 204; DB 10; Length 298; 76.0%; Pred. No. 1.1e-19; ative 2; Mismatches 10; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOSTWARE: PALENTIN VETSION 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 126, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION
APPLICANT: Busby, Robert
APPLICANT: Hecht, Peter
APPLICANT: Hecht, Peter
APPLICANT: Madden, Revin
APPLICANT: Madden, Revin
APPLICANT: Maxon, Mary
APPLICANT: Maxon, Mary
APPLICANT: Mille, Todd
APPLICANT: Molecula APPLICANT: Applicant: John
APPLICANT: Wille, Todd
APPLICANT: No. US20020128250Alman, Thea
                                                  Sequence 230, Application US/09801368
Patent No. US/0020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Hoth, Peter
APPLICANT: Madden, Kevin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Pencillium chrysogenum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Salama, Sofie
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sherman, Amir
Silva, Jeff
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 38; Conserv
                                   US-09-801-368-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-801-368-126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-801-368-230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
APPLICANT: Yanda, Yoji
TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDI
FILE REFERENCE: 00766.000043
CURRENT APPLICATION NUMBER: US/09/749,728B
CURRENT APPLICATION NUMBER: H11-372826
PRIOR APPLICATION NUMBER: H11-372826
PRIOR FILING DATE: 1999-12-28
PRIOR FILING DATE: 2000-02-28
PRIOR FILING DATE: 2000-02-28
PRIOR FILING DATE: 2000-11-02
PRIOR FILING DATE: 2000-11-02
NUMBER OF SEQ ID NOS: 80
SOFTWARE: PATENTIN VET.2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                   Score 201; DB 10; Length 730;
Pred. No. 7.7e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                              306 CFNCKTFKTPLWRRSPEGNTLCNACGLEQKLHGTMRPLSLKSDVIKKR 353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                     8; Indels
                                                                                                                                                                                                      1 CINCFIQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65.1%; Score 183; DB 10;
60.0%; Pred. No. 1.1e-16;
tive 10; Mismatches 10;
                                                                                                                                                     4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Maxon, Mary
Milne, Todd
No. US20020128250Alman, Thea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; Sequence 360, Application US/09801368; Patent No. US20020128250A1; GENERAL INFORMATION: APPLICANT: Busby, Robert; APPLICANT: Cali, Brian; APPLICANT: Holtch, Peter; APPLICANT: Holtch, Peter; APPLICANT: Madden, Kevin
                                                                                                                                                                                                                                                                                                                                                                           ; Sequence 11, Application US/09749728B
; Patent No. US20020142457A1
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-126
```

Gaps

ó

```
APPLICANT: Holtzman, Doug
APPLICANT: Holtzman, Doug
APPLICANT: Milna, G. Todd
APPLICANT: Milna, G. Todd
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Sherman, Amir
APPLICANT: Stary, Jeffry C.
APPLICANT: Truchemart, Josh
APPLICANT: Lixin
TITLE OF INVENTION: No. US20020182708Alel Regulators of Fungal Gene Expression
FILE REFRENCE: MIC-004
CURRENT APPLICATION NUMBER: US/10/029,180
CURRENT APPLICATION NUMBER: US 60/257,431
PRIOR APPLICATION NUMBER: US 60/257,431
PRIOR PILLING DATE: 2000-12-22
PRIOR FILLING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 138
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 68
LENGTH 359
                                                                                                                                                                                                                                                                                                                                                                Length 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 93, DB 9; Length 359;
Pred. No. 7.2e-05;
3; Mismatches 9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                            Indels
      TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
                                                                                                                                                                                                                                                                                                                                                                                                                                                      6 TQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 TITITLWRRNANGDPVCNACGLYYKLHNVNRPLTMKKEGIQTRNR 46
                                                                                                                                                                                                                                                                                                                                                              Ouery Match 56.2%; Score 158; DB 10; Best Local Similarity 62.2%; Pred. No. 5.3e-14; Matches 28; Conservative 7; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: Patentin Ver. 2.0
SOFTWARE: 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CTNCFTQTTPLWRRNPDG-QPLCNACGL-FLKL 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 124, Application US/10029180 Publication No. US20020182708A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-10-029-180-68
; Sequence 68, Application US/10029180
; Publication No. US20020182708A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; OTHER INFORMATION: fungal gene US-10-029-180-68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: Cali, Brian M.
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin T.
APPLICANT: Milna, G. Todd
APPLICANT: Sherman, Amir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 57.64
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT: Cali, Brian M.
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin 1
APPLICANT: Milna, G. Todd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sherman, Amir
                                                                                                                                                                                                                                                                                           ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                       US-09-925-300-1573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-10-029-180-124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                           TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: SUMMETS, Elic
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 69/487,558
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR PILING DATE: 1999-10-20
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEC ID NOS: 440
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                63.7%; Score 179; DB 10; Length 532; 58.0%; Pred. No. 4.4e-16; Live 7; Mismatches 14; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 549;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CTNCFTQFTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63.3%; Score 178; DB 10;
58.0%; Pred. No. 6.2e-16;
Live 8; Mismatches 13;
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR PLICATION NUMBER: US 60/160,587
PRIOR PLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Maxon, Mary
Milne, Todd
No. US20020128250Alman, Thea
Royer, John
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Sequence 1573, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 358, Application US/09801368 Patent No. US20020128250A1
                                                                                                                                                                                                                                                  ; ORGANISM: Penicillium chrysogenum US-09-801-368-360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM: Aspergillus nidulans US-09-801-368-358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 58.09
Matches 29; Conservative
                                                                                                                                                                                                                                                                                                                                                              29; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holtzman, Doug
Madden, Kevin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Salama, Sofie
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sherman, Amir
Silva, Jeff
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Craig Rosen, APPLICANT: Steve Ruben
                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-925-300-1573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
US-09-801-368-358
                                                                                                                                                                               SEQ ID NO 360
LENGTH: 532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
APPLICANT:
                                                                                                                                                                                                                               TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                            δ
```

ä

Gaps

;

```
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 499 CVSCHSSDSPCWR--PSWSPRKQDQLCNSCGLRYKKTHTRCLNDLCRKIPTKGEINIMKS 556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CINCFTQTTPLWRRNPDGQP-----LCNACGLFLK-----LHGVVRPLSLKTDV-IKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 6, Application US/09948018
Patent No. US20020150977A1
GENERAL INFORMATION:
APPLICANT: Theill et al
TITLE OF INVENTION: TWE RECEPTOR-LIKE MOLECULES AND USES THEREOF
FILE REPERENCE: 01017/37677
CURRENT PAPLICATION NUMBER: US 60/29048,018
PRIOR PRILING DATE: 2001-09-05
PRIOR PILING DATE: 2000-09-05
NUMBER OF SQ ID NOS: 45
SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 75; DB 10; Length 588;
Pred. No. 0.031;
9; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9 TPLWRR-----NPDGQPLCNACGLFLKLHGVVRP 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20.1%; Score 56.5; D
40.0%; Pred. No. 1.6;
tive 3; Mismatches
                                                                                                                                                                                                                                                                FILE REFERENCE: 109272.147

CURRENT APPLICATION NUMBER: US/09/801,368

CURRENT APPLICATION NUMBER: US/09/801,368

PRIOR APPLICATION NUMBER: US 09/487,558

PRIOR FILING DATE: 2000-01-19

PRIOR APPLICATION NUMBER: US 60/160,587

PRIOR FILING DATE: 1999-10-20

NUMBER OF SEQ ID NOS: 440

SOFTWARE: PatentIn version 3.0
                                                                                                                 No. US20020128250Alman, Thea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Saccharomyces cerevisiae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Sequence 2, Application US/09948018
; Patent No. US20020150977A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.7%;
32.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 40.0
Matches 14; Conservative
                                                                                                                                                                                   Amir
                                                                                                                                                            Salama, Sofie
                                                    Madden, Kevin
           Hecht, Peter
                                                                          Maxon, Mary
Milne, Todd
                                                                                                                                          Royer, John
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Homo sapiens
                                                                                                                                                                                                       Silva, Jeff
                                                                                                                                                                                     Sherman,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JS-09-801-368-46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 46
LENGTH: 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 6
LENGTH: 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
US-09-948-018-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 12
US-09-948-018-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-948-018-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     557 N 557
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49 N 49
                                                                                                                                                                              APPLICANT:
APPLICANT:
                                                                                                                 APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Пр
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT: Milna, G. Todd
APPLICANT: Sherman, Amir
APPLICANT: Shang, Lixin
APPLICANT: Zhang, Lixin
APPLICANT: Zhang, Lixin
TITLE OF INVENTION: NO. US20020182708Alel Regulators of Fungal Gene Expression
FILE REFERENCE: MIC-004
CURRENT APPLICATION NUMBER: US/10/029,180
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 60/257,431
PRIOR APPLICATION NUMBER: US 60/257,431
PRIOR FILING DATE: 2000-12-22
NUMBER OF SEQ ID NOS: 138
SSCTWARF: FastSEQ for Windows Version 4.0
SSCTWARF: FastSEQ for Windows Version 4.0
APPLICANT: Silva, Jeffry C.

APPLICANT: Trucheart, Josh
APPLICANT: Trucheart, Josh
TITLE OF INVENTION: No. US20020182708Alel Regulators of Fungal Gene Expression
FILE REFERENCE: MIC-004
CURRENT APPLICATION NUMBER: US/10/029,180
CURRENT FILING DATE: 2001-12-22
PRIOR APPLICATION NUMBER: US 60/257,431
PRIOR FILING DATE: 1000-12-22
NUMBER OF SEQ ID NOS: 138
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5;
                                                                                                                                                                                                                                                                                                                                                                                                                          Length 440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 451;
                                                                                                                                                                                                                                                                                                                                                                                                                        Score 93; DB 9; Length 440
Pred. No. 9.1e-05;
3; Mismatches 9; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 93; DB 9; Le
Pred. No. 9.3e-05;
3; Mismatches 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CINCFIQTTPLWRRNPDG-QPLCNACGL-FLKL 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 CTNCFTQTTPLWRRNPDG-QPLCNACGL-FLKL 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 123, Application US/10029180
Publication No. US20020182708A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 46, Application US/09801368; Patent No. US20020128250al; GENERAL INFORMATION: APPLICANT: Busby, Robert appLICANT: Call, Brian
                                                                                                                                                                                                                                                                                                                                                      ; OTHER INFORMATION: fungal gene US-10-029-180-124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; OTHER INFORMATION: fungal gene US-10-029-180-123
                                                                                                                                                                                                                                                                                         TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 33.1%;
Best Local Similarity 57.6%;
Matches 19; Conservative 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 33.1%;
Best Local Similarity 57.6%;
Matches 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Cali, Brian M.
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin T.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-10-029-180-123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-801-368-46
                                                                                                                                                                                                                                           SEQ ID NO 124
LENGTH: 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURE:
                                                                                                                                                                                                                                                                                                                                      FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
```

```
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
US-09-948-018-36
i Sequence 36, Application US/09948018
j Patent No. US20020150977a1
i GENERAL INFORMATION:
i APPLICANT: Theill et al
i TITLE OF INVENTION: TWE RECEPTOR-LIKE MOLECULES AND USES THEREOF
i FILE REFERENCE: 01017/37677
CURRENT APPLICATION NUMBER: US/09/948,018
i CURRENT FILING DATE: 2001-09-05
i PRIOR RPLICATION NUMBER: US 60/230,191
i PRIOR FILING DATE: 2000-09-05
i NUMBER OF SEQ ID NOS: 45
i SOFTWARE: Patentin version 3.1
           APPLICANT: THE LIGHT AT THE RECEPTOR-LIKE MOLECULES AND USES THEREOF FILE REPERENCE: 01017/3762 THE RECEPTOR-LIKE MOLECULES AND USES THEREOF CURRENT APPLICATION UNDBER: US/09/948,018
CURRENT APPLICATION NUMBER: US 60/230,191
PRIOR APPLICATION NUMBER: US 60/230,191
PRIOR APPLICATION NUMBER: US 60/230,191
PRIOR FILING DATE: 2000-09-05
SOFTWARE: PALENTH VERSION 3.1
SEQ ID NO 2
LENGTH: 355
                                                                                                                                                                                                                                                                                                                                                                                                  11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11;
                                                                                                                                                                                                                                                                                                                                                    DB 10; Length 355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 10; Length 380;
                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                    Score 56.5; D
Pred. No. 4.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                       9 TPLWRR-----NPDGQPLCNACGLFLKLHGVVRP 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 56.5; DI
Pred. No. 5.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 20.1%; Score 56.5; D
Best Local Similarity 40.0%; Pred. No. 4.9;
Matches 14; Conservative 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9 TPLWRR-----NPDGQPLCNACGLFLKLHGVVRP 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 5832, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: ANDO, SEIKO
APPLICANT: ANDO, SEIKO
APPLICANT: ANDO, SEIKO
APPLICANT: YOKOI, HARNHKO
APPLICANT: YOKOI, HARNHKO
APPLICANT: YOKOI, HARNHKO
APPLICANT: TATEISHI, NAOKO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 20.1%;
Best Local Similarity 40.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SENOH, AKIHIRO
IKEDA, MASATO
                                                                                                                                                                                                                                                          TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-018-36
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
US-09-738-626-5832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 36
LENGTH: 380
                                                                                                                                                                                                                                                                                                      US-09-948-018-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
```

|   |   |  |                                |   | ·             |
|---|---|--|--------------------------------|---|---------------|
|   |   |  |                                |   |               |
|   |   |  |                                |   |               |
|   |   |  |                                |   |               |
|   |   |  |                                |   |               |
|   |   |  |                                |   |               |
|   |   |  |                                | ^ |               |
|   | • |  |                                |   |               |
| , | - | The second section of the second section of the second section of the second section s | uni andreagan angan angan dire |   | andersonador. |
|   |   |  |                                |   | <i>;</i>      |

```
January 3, 2003, 18:48:02; Search time 44.5 Seconds (without alignments) 149.720 Million cell updates/sec
                                                                                                                                                                                                                       281
1 CTNCFTQTTPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR 50
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                    Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                   908470 seqs, 133250620 residues
                                                                              OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
                                                                                                                                                                                                        US-09-725-010-3
                                                                                                                                                                                                                         Perfect score:
                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database :
                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                               Sequence:
                                                                                                                     Run on:
```

A\_Geneseq\_101002:\*

1 /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
10: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1989.DAT:\*
11: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
12: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
13: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
14: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
15: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
16: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
17: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
18: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
19: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
10: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
11: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
12: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
13: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
14: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
15: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
16: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
17: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
18: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
19: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
21: /SIDS2/gcgdata/geneseqg/geneseqp-embl/AA1999.DAT:\*
22: /SIDS2/gcgdata/geneseqg/geneseqgp-embl/AA1999.DAT:\*
23: /SIDS2/gcgdata/geneseqg/geneseqgp-embl/AA1999.DAT:\*
23: /SIDS2/gcgdata/geneseqg/geneseqgp-embl/AA1999.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Description                     | NIT2 DNA binding d<br>Amino acid sequenc<br>AreA activator pro<br>Aspergillus oryzae<br>GAT99 DNA binding<br>Drosophila melanog<br>Human colon cancer<br>Human ovarian anti<br>Human GATA-6 trans<br>Human GATA-6 trans |
|---------------------------------|---|
| SUMMARIES                       | AABB2413<br>AAWB1079<br>AARB8408<br>AAW31630<br>ABB2414<br>ABB71271<br>AAG74577<br>AAG74577<br>AAW56703<br>AAW56703   |
| DB 1                            | 22222222222222222222222222222222222222  |
| engt                            | 8853<br>866<br>866<br>866<br>866<br>870<br>870<br>870<br>870<br>844<br>844  |
| %<br>Query<br>e Match Length DB | 100.0<br>98.6<br>98.6<br>91.1<br>67.6<br>66.9<br>66.9   |
| Score                           | 281<br>277<br>277<br>277<br>277<br>256<br>190<br>188<br>188<br>188  |
| Result<br>No.                   | 10<br>10<br>10  |

| Human GATA4 protei<br>Heart muscle cell<br>Drosophila melanog<br>Mouse GATA-3 trans<br>Transgenic animal<br>Pain regulated pro | Jeyurateu 3 amino ac 3 phila mela 5 protein s 5 cancer p 6 breast ca | inger<br>lopsis<br>lopsis<br>lopsis<br>ORFX<br>lopsis                | is thal<br>is thal<br>is thal<br>is tran<br>is thal<br>is thal<br>is thal<br>is thal<br>is thal                      |
|--|--|--|--|
| AAGG4849<br>AAGG4849<br>ABB64550<br>AAU79801<br>ABB85021<br>ABB85021   | AAB9251<br>AAB9251<br>AAB9251<br>AAB9019<br>AAB0019<br>AAB4055       |  |  |
| 0000000  | 2444<br>1128<br>1128<br>2281<br>2281<br>294<br>294                   | 273 222<br>273 222<br>271 21<br>270 21<br>276 21<br>276 21<br>276 21 | 294 21<br>295 21<br>295 21<br>163 23<br>264 23<br>285 21<br>287 21<br>287 21<br>295 21<br>295 21<br>295 21<br>295 21 |
| طططشة والم   | 500000000000000000000000000000000000000                              |  | NNNN 44444444<br>© © © O O O O O O O O O O O O O O O O O   |
| 8887777  | 178<br>178<br>176<br>139<br>139<br>139<br>139<br>128                 | 08.<br>08.<br>08.<br>10<br>10<br>00.                                 | 1000.5<br>1000.5<br>1000.5<br>966.5<br>966.5<br>966.5<br>966.5   |
| 11111111111111111111111111111111111111   | 222<br>222<br>243<br>243   | 3335555<br>3335555<br>3335555555555555555555                         | 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9  |

## ALIGNMENTS

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR88408
                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     XXXX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SO
                                  The present sequence is that of the NIT2 DNA binding domain of Neurospora crassa. The NIT2 protein is a global positive-acting transcription factor of nitrogen structural genes when preferred N-sources are lacking. The promoter region (see AAP9041) of the Candida albicans hyphal-specific HWP1 gene contains numerous binding sites for NIT2. GAT99 (see AAB8414), showing 92% homology to NIT2, was identified in C. albicans, suggesting that regulation of HWP1 expression may be tied to nitrogen regulatory events. The HWP1 gene encodes an adhesin required for the pathogenesis of candidiasis. The invention provides a method for interfering with the expression of hyphal-specific genes in a fungus resulting in inhibition of cell growth. This involves interfering with the transcription of the hyphal-specific gene anipulating the binding of a DNA binding protein to a cis-regulatory element. The fungus may be a non-pathogenic or a cis-regulatory element. The fungus may be a non-pathogenic or a pathogenic fungus, such as C. albicans, and the hyphal-specific gene may be HWP1, in which the NTP binding sites act as the cis-regulatory elements. The method provides a means of treating infection of the means of treating infection of the means.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                     elements. The method provides a means of treating intection of mammalian hosts such as immunocompromised or immunosuppressed humans, including those having AIDS or undergoing transplantation or anti-cancer therapy, burns patients, patients with diabetic ketoacidosis, and patients in which the normal microbial flora has been disrupted because of disease, trauma or chemical, radiation or other immunosuppressive prophylaxis. Also provided is a method for characterising genes under control of a DNA binding protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New koji mold capable of expressing two times more endo- and exo-peptidases than wild-type strain Aspergillus oryzae CNCM I-1882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Amino acid sequence of the Aspergillus oryzae areA protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               areA gene; peptidase; Koji mold; L-glutamine; prolidase; prolyl-dipeptidyl-peptidase. \angle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 281; DB 22;
100.0%; Pred. No. 1.7e-29;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAW81079 standard; Protein; 853 AA.
               Claim 31; Page 80; 95pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Affolter M, Van Den Broek P;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-EP02785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97EP-0111378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (NEST ) SOC PROD NESTLE SA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-MAY-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1999-120892/10.
N-PSDB; AAV68650.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Aspergillus oryzae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                       50 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09902691-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            05-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAW81079;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAW81079
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
```

```
Gaps
useful for hydrolyzing protein-containing materials, and producing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        oryzae IFO4177
                                                                                                                                            This is the amino acid sequence of the area protein of Aspergillus oryzae used in the method of the invention involving the doubled expression of peptidases in a new Koji mold. The Koji mold is useful for hydrolyzing protein-containing materials, and for over-producing proteolytic enzymes. When hydrolyzing protein-containing materials (Containing & Ma. L-glutamine), the mold may be used in combination with an enzyme and/or a microorganism providing a prolidase activity. Additionally, the mold is useful in fermentation processes for creating food products. The koji mold is used to produce high levels of endopeptidases and exopeptidases, including enhanced prolyl-dispeptiday-peptidase activity, which may be useful for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The AreA activator protein (AAR88408) of Aspergillus oryzae Ir controls nitrogen catabolism and influences the production of extracellular proteases. It is the product of the areA gene (AAT09946). Inactivation of the activator, by deletion or antisense inhibition of the gene, yields protease-free Aspergillus cells, e.g. strain ToC913, suitable as hosts for produc of protease-susceptible heterologous proteins such as enzymes and industrial or therapeutic proteins.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Fungus with a modified areA gene - useful to produce proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98.6%; Score 277; DB 20; Length 853; 98.0%; Pred. No. 1.3e-27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFIQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                    food products with improved organoleptic properties
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         · 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      susceptible to degradation by protease(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Area activator; protease-free; host cell.
                                                                                                 Claim 11; Pages 28-30; 53pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 23; Page 23-25; 42pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    releasing L-glutamine from peptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR88408 standard; Protein; 866 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             spergillus oryzae strain IFO4177.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94DK-0000717.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            95WO-DK00254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Christensen T, Hynes MJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (NOVO ) NOVO-NORDISK AS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Area activator protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1996-058422/06.
N-PSDB; AAT09946.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       853 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-JUN-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JUN-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9535385-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-JUN-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-DEC-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR88408;
```

ö

AAB82414 standard; Protein; 50 AA.

RESULT 5 AAB82414

ö

Gaps

AAB82414;

```
This protein comprises the AreA regulator protein of Aspergillus oryzae. The amino acid sequence was deduced from the isolated areA gene (see AAT97394). The invention also relates to new fungi in which: (a) the areA gene has been modified by recombinant DNA methods so that it cannot express a functional AreA activator; and (2) the extracellular protease pepc and/or pepg genes are inactivated so that functional exarmes (see AAM31628-29) are not produced. The new fungi are used for the recombinant production
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of industrial or therapeutic peptides and proteins, particularly enzymes (e.g. proteases, lipase, cutinase, cellulase or chymosin), insulin, growth hormone, glucagon, somatostatin, interferon, platelet derived growth factor, factor VII, factor VIII, urokinase,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tissue plasminogem activator, erythropoietin or thrombopoietin. The new fungi do not produce protease and so provide higher yields of recombinant protein because of reduced proteolytic degradation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        extracellular protease industrial or therapeutic
                                                                                                                                                                                                                                                                                                                                                                AreA regulator protein; PepE; PepC; protease; gene replacement; antisense; gene inactivation; protease-free host cell; recombinant protein.
                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 277; DB 18; Length 866;
Pred. No. 1.3e-27;
1; Mismatches 0; Indels 0
                                98.6%; Score 277; DB 17; Length 866; 98.0%; Pred. No. 1.3e-27;
                                                                                                                      664 CINCFTQTTPLWRRNPEGQPLCNACGLFLKLHGYVRPLSLKTDVIKKRNR 713
                                                                                                      1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New fungus lacking functional AreA and extracellular pr
gene(s) - for high yield expression of industrial or th
proteins, also new protease(s) from Aspergillus oryzae
                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                 Aspergillus oryzae AreA regulator protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 3; Page 46-48; 80pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                       Aspergillus oryzae strain IFO 4178.
                                                                                                                                                                                                                           AAW31630 standard; Protein; 866
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               96WO-DK00528.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                95DK-0001428.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Christensen T, Lehmbeck J;
                                                                                                                                                                                                                                                                                                (first entry)
                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (NOVO ) NOVO-NORDISK AS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1997-341696/31.
N-PSDB; AAT97394.
                                                 Sest Local Similarity
866 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    866 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9722705-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16-DEC-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-DEC-1995;
                                                                                                                                                                                                                                                                                                11-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-JUN-1997
                                                                   46;
 Sednence
                                                                                                                                                                                                                                                             AAW31630;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                Query Match
                                                                     Matches
                                                                                                                                                                                          RESULT 4
                                                                                                                                                                                                          AAW31630
g
                                                                                                                                                                                                                             ŏ
                                                                                                                                     qq
```

```
The present sequence is that of the GAT99 (GAT-1 or GATA-like protein) of Candida albicans. The sequence shows 92% homology to protein) of Candida albicans. The sequence shows 92% homology to positive-acting transcription factor of nitrogen structural genes when preferred N-sources are lacking. The promoter region (see AAF90471) of the Candida albicans hyphal-specific GAT99 suggests that regulation of HWP1 expression may be tied to nitrogen regulatory events. The HWP1 gene encodes an adhesin regulated for the pathogenesis of candidiasis. The invention provides a method for interfering with the expression of hyphal-specific genes in a fungus resulting in inhibition of cell growth. This involves interfering with the transcription of the hyphal-specific gene manipulating the binding of a DNA binding protein to a cis-regulatory element. The fungus may be a non-pathogenic or a pathogenic fungus, such as C. ablicans, and the hyphal-specific gene may be HWP1, in which the NT2 binding sites act as the cis-regulatory elements. The method provides a means of treating infection of elements. The mannosuppressed humans, including those having AIDS or undergoing transplantation or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interfering with expression of hyphal-specific genes in fungus for inhibiting fungal cell growth involves interfering with transcription of hyphal-specific genes mediated by cis acting sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   been disrupted because of disease, trauma or chemical, radiation or other immunosuppressive prophylaxis. Also provided is a method for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             burns patients, patients with diabetic
                                                                                                                            Hyphal wall protein; HWP1; hyphal-specific gene; infection; candidiasis; therapy; antifungal; fungicide; transcription factor; GAT99; GAT-1; GATA-like protein; DNA binding protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      other immunosuppressive prophylaxis. Also provided is a met
characterising genes under control of a DNA binding protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 256; DB 22;
Pred. No. 3.2e-26;
1; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 32; Page 80; 95pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91.1%;
92.0%;
                                                                                                                                                                                                                                                                                                                     29-NOV-2000; 2000WO-US32464.
                                                                                                                                                                                                                                                                                                                                                            99US-0167672.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and patients
                                                     06-AUG-2001 (first entry)
                                                                                        GAT99 DNA binding domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   anti-cancer therapy,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-367698/38.
                                                                                                                                                                                                                                                                                                                                                                                                (SUND/) SUNDSTROM P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                50 AA;
                                                                                                                                                                                                       Candida albicans.
                                                                                                                                                                                                                                            WO200138550-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ketoacidosis,
                                                                                                                                                                                                                                                                                                                                                         29-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                       Sundstrom P;
                                                                                                                                                                                                                                                                                  31-MAY-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
```

;

1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50 

à

; 0

Gaps

0;

1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50

98.6%; 98.0%;

Conservative

Local Similarity nes 49; Conserv

Best Loca Matches

δŏ

Query Match

ABB71271;

RESULT

g

```
cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in dene therapy and vacche production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing inactive proteins or to supplement the patients own production of P. Additionally, N may be used to produce the colon cancer-associated Ps, by inserting the nucleic acids into a host cell and culturing the cell coxpress the proteins. N and P can be used in the prevention, diagnosis and AAB77789 represent sequences used in the exemplification of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour; ovarian cancer; breast cancer; tumour; reproductive system disorder; infertility; pregnancy disorder; anovulation; polycystic ovary syndrome; PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection; inflammatory condition; immune disorder; blood disorder; cardiovascular disorder; respiratory disorder; neurological disorder;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4277~\text{human} colon cancer-associated polypeptides, diagnosing and/or treating colorectal cancers -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma; chromosome 18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 188; DB 22;
Pred. No. 1.8e-16;
9; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human ovarian antigen HUSGX12, SEQ ID NO:3017.
                                                                                                                                                                                                                                                                                                                                                                  Rosen CA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 11; Page 6989-6990; 9803pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ABP41885 standard; Protein; 279 AA.
                                                                                                                                                                                                                                                                                                                                                                Birse CE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66.9%;
62.0%;
                                                                                                                                                                                                          28-SEP-2000; 2000WO-US26524.
                                                                                                                                                                                                                                                     99US-0157137.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nucleic acids encoding useful for preventing,
                                                                                                                                                                                                                                                                                                                                                                Barash SC,
                                                                                                                                                                                                                                                                                                                    (HUMA-) HUMAN GENOME
                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-235357/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       279 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAH33982
                                                                                                                      WO200122920-A2
                                                                             Homo sapiens.
                                                                                                                                                                                                                                                     29-SEP-1999;
                                                                                                                                                                                                                                                                            03-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-AUG-2002
                                                                                                                                                                 05-APR-2001
                                                                                                                                                                                                                                                                                                                                                                Ruben SM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABP41885;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABP41885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL016176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                         Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; SEQ ID NO 40605; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 950;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                Drosophila melanogaster polypeptide SEQ ID NO 40605.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human colon cancer antigen protein SEQ ID NO:5341.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 190; DB 22;
Pred. No. 3.8e-16;
7; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Myers EW;
                                                                                                                      Š
                                                                                                               ABB71271 standard; Protein; 950
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAG74577 standard; Protein; 279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67.68;
64.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23-MAR-2000; 2000US-191637P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11-JUL-2000; 2000US-0614150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       03-SEP-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                           Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Venter JC, Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-656860/75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  950 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; ABL15374
                                                                                                                                                                                                                                                                                                                                                                                                    WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interactions -
                                                                                                                                                                                                                                                                                                                  pharmaceutical
                                                                                                                                                                                                   26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                               27-SEP-2001
```

.. 0

Gaps

.; 0

32;

Matches

RESULT 7

XEXEXEX

g à

Seguence Query Match

```
ABP43228) and to CDNAs encoding them (ABG94131-ABG96305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen to the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen to opport of ovarian antigen polynucleotides against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treathny, prognosing or preventing various ovary and/or breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and variantis), immune disorders (e.g., congenital and acquired immunodeficiencies, autofimmune ophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders (e.g., anaemia), cardiovascular disorders explaintinestinal disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        further be used for gene therapy, chromosome mapping, in the diadrification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           invention relates to 2175 novel human ovarian antigens (ABP41054-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                   gene therapy, chromosome mapping; forensic analysis; antibody preparation; cytostatic; immunomodulatory; neuroprotective; antiinflammatory; gynaecological; reproductive.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Isolated nucleic acid molecules encoding novel ovarian polypeptides,
gastrointestinal disorder; urinary system disorder; drug screening;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 188; DB 23; Length 279;
Pred. No. 1.8e-16;
9; Mismatches 10; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 11; SEQ ID No 3017; 2922pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     66.9%;
62.0%;
                                                                                                                                                                                                                                                                                                                                                          (HUMA-) HUMAN GENOME SCI INC
                                                                                                                                                                                                                                                               07-JUN-2001; 2001WO-US18569.
                                                                                                                                                                                                                                                                                                           07-JUN-2000; 2000US-209467P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              neurological diseases
                                                                                                                                                                                                                                                                                                                                                                                                        Rosen CA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2002-147878/19.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      279 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; ABQ54962
                                                                                                                                                               WO200200677-A1
                                                                                                                     Homo sapiens.
                                                                                                                                                                                                                03-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                        Birse CE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
  δ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Use of GATA-6 transcription factor - for inhibiting vascular smooth muscle cell proliferation, e.g. for treatment of arteriosclerosis or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This polypeptide comprises human GATA-6 transcription factor, a protein that is expressed in differentiated vascular smooth muscle cells (vSMC) and which prevents mitogen stimulated proliferation of such cells. The amino acid sequence was deduced from human GATA-6 CDNA (see AAV28411). It contains 2 zinc-finger domains followed by a highly conserved region containing a high frequency of basic amino acids. A claimed method for treating a subject diagnosed as having a condition associated with excessive VSMC proliferation comprises administering to the subject an isolated GATA-6 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or nucleic acid that prevents or reduces excessive VSMC proliferation in vivo. This method is used to treat an arteriosclerotic condition such as atherosclerosis and port interventional restenosis. GATA-6 polypeptides can also be used
                                                                                                      GATA-6; transcription factor; vascular smooth muscle cell; VSMC;
                                                                                                                     cell proliferation; atherosclerosis; restensis; tumour; therapy;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 449;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               298 CANCHTTTTLWRRNAEGEPVCNACGLYMKLHGVPRPLAMKKEGIQTRKR 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 CINCFIQITIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 188; DB 19; Length 4
Pred. No. 3e-16;
9; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                      (SELI-) ST ELIZABETH'S MEDICAL CENT BOSTON INC.
                                                                                                                                                                                               Location/Qualifiers
244..306
/note= "zinc-finger domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 13; Page 48-50; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB99920 standard; Protein; 442 AA
                                                                         Human GATA-6 transcription factor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  66.9%;
62.0%;
                                                                                                                                                                                                                                                                                                                                                       96US-0025574.
                                                                                                                                                                                                                                                                                                                         97WO-US14832
                                         28-AUG-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26-SEP-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 66.99
Best Local Similarity 62.0
Matches 31; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in drug screening assays.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1998-193638/17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         449 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAV28411
                                                                                                                                                                   Homo sapiens.
                                                                                                                                                                                                                                                                                                                         22-AUG-1997;
                                                                                                                                                                                                                                                                                                                                                       06-SEP-1996;
                                                                                                                                                                                                                                                            WO9810090-A1
                                                                                                                                                                                                                                                                                            12-MAR-1998.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Seguence
              AAW56703;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB99920;
                                                                                                                                                                                                                                                                                                                                                                                                                      Walsh K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tumours
                                                                                                                                      human.
                                                                                                                                                                                                                Domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB99920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DXAXX
```

Human GATA4 protein sequence SEQ ID NO:11.

AAW56703 standard; Protein; 449 AA

AAW56703 ID AAW5 RESULT 9

Op

ö

```
Angiogenesis; cardiant; cell differentiating agent; bone marrow;
                                           Homo sapiens
                                                                                                                                                   28-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-FEB-2000;
02-NOV-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-SEP-2001
                                                                                              05-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         05-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Umezawa A,
                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yamada Y;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAG64849;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
AAG64849
  QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           umbilical blood cells which are capable of differentiating into cardiomyocytes. Also described are: (1) cardiomyocytes produced by the differentiation of the cells; (2) a method for carrying out the differentiation of the cells; (2) a method for carrying out the differentiation into cardiomyocytes, regulated by a promotional and/or inhibitory factor; (3) a method for the differentiation of the cells into cell types other than cardiomyocytes; (4) drug compositions promoting the formation of heart muscle and regeneration of heart tissue which contain the cells; (5) a method for the production of antibodies which recognise the cells; (5) a method for screening factors which promote the proliferation of the cells; (7) a method for immortalising the treatment of heart disease which contain the immortalising the treatment of heart disease which contain the immortalising of the cells by expressing telomerase in them; (8) drug compositions for the treatment of heart disease which contain the immortalising of 19 cell-free supernatual from the culture of the cells and its use in promoting their differentiation into cardiomyocytes. The cells are used in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction and in the study of cardiomyocyte differentiation. AAH44351 to AAH44409 and AAB99915 to AAB99935 represent sequences used in the exemplification of the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                     Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
         Differentiation; heart muscle cell; cytokine; transcription factor; proliferation; surface antigen; heart disease; cardiomyocyte; bone marrow; umbilical blood cell; heart muscle degeneration;
                                                                                                                                                                                                                                                                                                                                                                                                                       present invention describes cells originating in bone marrow or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                             ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65.1%; Score 183; DB 22; Length 442; 60.0%; Pred. No. 1.4e-15;
                                                                                                                                                                                                                                                            Sakurada K, Gojo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10; Mismatches
                                                                                                                                                                                                                                                            Ogawa S,
                                                                                                                                                                                                                                                                                                                                                                                         Claim 41; Page 98-99; 18/pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAG64319 standard; Protein; 442 AA.
                                                                                                                                                                                                                                                         Fukuda K,
                                                                                                                                                                                                                               (KYOW ) KYOWA HAKKO KOGYO KK.
                                                                                                                                                                                       28-DEC-1999; 99JP-0372826.
28-FEB-2000; 2000WO-JP01148.
                                                                                                                                                             02-NOV-2000; 2000WO-JP07741.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24-SEP-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                     myocardial infarction.
                                                                                                                                                                                                                                                          Hata J,
                                                                                                                                                                                                                                                                                             WPI; 2001-425655/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human GATA4 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           442 AA;
                                                                                                                                                                                                                                                                                                              N-PSDB; AAH44356.
                                                                                                        WO200148150-A1.
                                                                                Homo sapiens.
                                                                                                                                  05-JUL-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30;
                                                                                                                                                                                                                                                            Umezawa A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                       Yamada Y;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAG64319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
AAG64319
XEXEXEX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
The present invention relates to cells isolated from bone marrow, which are capable of at least differentiating into heart muscle cells. The cells are applicable as remedies for various heart diseases particularly with damaged heart muscle accompanying degeneration. The present sequence was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New adult bone marrow-originated cells capable of differentiating into
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 heart muscle cells, applicable as remedies for various heart diseases particularly with damaged heart muscle accompanying degeneration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Heart muscle cell differentiation related protein SEQ ID NO: 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Heart muscle cell; human; cell differentiation; heart disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 442;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     271 CANCQTTTTLWRRNAEGEPVCNACGLYMKLHGVPRPLAMRKEGIQTRKR 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             90,09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                       Sakurada K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sakurada K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 22;
heart muscle cell; heart disease; human; GATA4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No. 1.4e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                          Ogawa S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ogawa S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65.1%; Score 183; 60.0%; Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 23; Pages 74-76; 158pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAG64849 standard; Protein; 442 AA.
                                                                                                                                                                                                                                                                                                                                                                                     Fukuda K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Fukuda K,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99JP-0372826.
2000WO-JP01148.
2000WO-JP07741.
                                                                                                                                                                                                                                                                                                                                 (KYOW ) KYOWA HAKKO KOGYO KK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (KYOW ) KYOWA HAKKO KOGYO KK.
                                                                                                                                                                                                                     28-FEB-2000; 2000WO-JP01148.
                                                                                                                                                                                                                                                                           99JP-0372826.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27-DEC-2000; 2000WO-JP09323.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 60.0%
Matches 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                       Umezawa A, Hata J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hata J,
                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-418252/44.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2001-425656/45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       442 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-PSDB; AAH49591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200148151-A1.
                                                                                                         WO200148149-A1
```

ó

Length 486;

486 AA;

õ g

```
at ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAU97801;
                                    Sequence
                                                                                                                                                                                                                                                                                                            AAU79800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAU97801
ID AAU9
XX
AC AAU9
                                                                                                                                                                                                                                      RESULT 14
                                                                                                                                                                                                                                                       AAU79800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                           δ
 SXS
                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13
                                                                                                                                                marrow
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                      and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
                                  Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of
                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                            The present invention provides cells originating in the human bone m or umbilical blood cells which are capable of differentiating into cardiomyocytes. These cells are useful in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction, the study of cardiomyocyte differentiation. The present sequence is a protein described in the exemplification of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; SEQ ID NO 20442; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                         271 CANCQTTTTTLWRRNAEGEPVCNACGLYMKLHGVPRPLAMRKEGIQTRKR 320
                                                                                                                                                                                                                                                                                                                                                                                     1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        discloses genomic DNA sequences (ABL16176-ABL30511), exsequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Drosophila melanogaster polypeptide SEQ ID NO 20442.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Myers EW;
                                                                                                            Claim 36; Page 104-106; 183pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABB64550 standard; Protein; 486 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (ABB57737-ABB72072).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Venter JC, Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2001-656860/75
                                                                                                                                                                                                                                                                                                                               Local Similarity
nes 30; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                            442 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; ABL08653
 N-PSDB; AAH48225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200171042-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Drosophila;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27-SEP-2001
                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABB64550
                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
ABB64550
```

```
ö
                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                expressing GATA-3 in a nonhuman mammalian cell, and which contains a DNA encoding ggata-3 or its mutant gene and is recombined together with a promoter/enhancer sequence so that it may be expressed at a specified site. The compound is useful as a preventive and/or treating agent for autoimmune diseases and allergic diseases. The recombinant vector may be used to generate a transgenic mouse highly expressing mouse GATA-3. The present sequence represents mouse GATA-3 transcription
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Transgenic animal selectively expressing GATA-3, a recombinant vector selectively expressing GATA-3 of a nonhuman mammalian cell, and a preventive and/or treating agent for immune diseases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present invention relates to a recombinant vector for selectively
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mouse, GATA-3 transcription factor; ggata-3; autoimmune disease; allergic disease; transgenic; immunosuppressive; antiallergic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                              ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 443;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 178; DB 23;
Pred. No. 6.1e-15;
Score 180; DB 22,
Pred. No. 3.7e-15,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8; Mismatches
                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 5-7; 8pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAU97801 standard; Protein; 443 AA.
                                                                                                                                                                                                                                                                             AA
                                                                                                                                                                                                                                                                                                                                                                                                                            Mouse GATA-3 transcription factor.
                                                                                                                                                                                                                                                                        AAU79800 standard; Protein; 443
                                              7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63.3%;
60.0%;
64.1%;
62.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      04-OCT-2000; 2000JP-0305252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       04-OCT-2000; 2000JP-0305252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SUMU ) SUMITOMO SEIYAKU KK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2002-447980/48.
N-PSDB; ABK50690.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
nes 30; Conserv
                            Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    443 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JP2002112781-A.
                                                                                                                                                                                                                                                                                                                                                                           13-AUG-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16-APR-2002
                                                   31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sednence
```

```
The invention describes a transgenic non-human mammal expressing transcription factor GATA-3 or its variant. The transgenic animal is useful for screening for agents for prevention and/or treatment of allergic diseases and autoimmune diseases. Prepared transgenic non-human mammals are useful for screening of treatment agent of allergic diseases and autoimmune diseases. This is the amino acid sequence of a protein associated with the transgenic animal expressing GATA-3 described in the
                                                                                                                                                                                                                                                                                                                                                                       Transgenic non-human mammals expressing GATA-3 or its variant prepared by mating a mouse with introduced a transfer factor GATA-3 gene and a mouse with introduced egg white albumin recognising T-cell receptor
                                                                              Transgenic animal; transcription factor; GATA-3; allergic disease; autoimmune disease.
                                              Transgenic animal expressing GATA-3 associated protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Page 6-7; 9pp; Japanese.
                                                                                                                                                                                                                          04-OCT-2000; 2000JP-0305253.
                                                                                                                                                                                                                                                       04-OCT-2000; 2000JP-0305253.
                                                                                                                                                                                                                                                                                        (SUMU ) SUMITOMO SEIYAKU KK
               13-AUG-2002 (first entry)
                                                                                                                                                                                                                                                                                                                    WPI; 2002-447972/48.
N-PSDB; ABK51619.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 443 AA;
                                                                                                                                                           JP2002112666-A.
                                                                                                                                                                                           16-APR-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seguence
                                                                                                                                Mus sp.
```

Search completed: January 3, 2003, 19:47:49 Job time : 45.5 secs

QY

0;

Gaps

; 0

63.3%; Score 178; DB 23; Length 443; 60.0%; Pred. No. 6.1e-15; Live 8; Mismatches 12; Indels (

Conservative

30;

Matches

Query Match Best Local Similarity

```
January . 3, 2003, 15:22:02; Search time 2756 Seconds (without alignments) 15522.904 Million cell updates/sec
                                                                                                                                                        US-09-725-010-1
1470
1 ggatctttctttttcatttc......cttttagtttcgtcaatatg 1470
                                                                                                                                                                                                                                                                                                        4109280
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                       2054640 segs, 14551402878 residues
                                                          nucleic search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                         IDENTITY_NUC Gapop 10.0 , Gapext 1.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  em_htgo_other:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_htg_inv:*
em_htg_other:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_htg_pln:
em_htg_rod:*
em_htg_rod:*
em_htg_rur:*
em_sy:*
em_htgo_hum:*
em_htgo_hum:*
                                                                                                                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_htg_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_htg_mus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                              9b_ba:*
9b_htg:*
9b_nin:*
9b_ow:*
9b_pox:*
9b_ph:*
9b_ph:*
9b_pr:*
9b_pr:*
9b_ri:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em_on:*
em_or:*
em_ov:*
em_pat:*
em_ph:*
em_pl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_sts:*
em_un:*
em_vi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                GenEmbl:*
                                                                                                                                                          Title:
Perfect score:
                                                                                                                                                                                                                       Scoring table:
                                                             OM nucleic -
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database :
                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                       Searched:
                                                                                           Run on:
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| scription                   | ! . <u></u>   | Plasmodi<br>Plasmodi<br>Plasmodi         | 'a'   | ACOUSTON TRANSMORTU<br>ACIT5598 Dictyoste<br>AL034559 Plasmodiu<br>ACIT7072 Dictyoste<br>AX345328 Sequence<br>AF211124 Carsonell<br>ACO11593 Homo sapi<br>ACO105994 Homo sapi |   | AC117072 Dictyoste AC009435 Homo sapi AXX46059 Sequence AC007158 Homo sapi AC007158 Homo sapi AXX86935 Sequence AC106883 Homo sapi AC107032 Homo sapi AC11559 Dictyoste AC115692 Dictyoste AC116962 Dictyoste AC116965 Dictyoste AC116965 Dictyoste AC116965 Dictyoste AC116965 Dictyoste AC116965 Dictyoste AC116965 Homo sapi AC106856 Homo sapi AC106856 Homo sapi AX083744 Sequence AX347067 Sequence | linear PAT 22-JUN-2001   | Saccharomycotina; Saccharomycetes;<br>Saccharomycetales; Candida.<br>ssion of hyphal-specific genes                                       |
|-----------------------------|---|--|---|---|---|---|--|---|
| SUMM)                       | 6 AX15464<br>8 CAU6420<br>8 AF00197<br>8 CAC35A5      | 3 PFMAL3P<br>2 AC00550<br>2 AC00415      | 2 ACLIDOS<br>3 AEO0139<br>3 PFMAL3P<br>6 AX45706<br>2 ACO0550 | 70m0v=m00   | ATT28A8<br>8 ATT28A8<br>3 AE00147<br>6 AX34660<br>2 AC11726<br>9 HSA3237<br>3 AE00139 | 2 AC117072<br>9 AX3009435<br>1 6 AX3009435<br>2 AC007158<br>6 AC206935<br>6 AC206935<br>9 AC107032<br>2 AC117073<br>2 AC116962<br>2 AC116962<br>9 AC106965<br>9 AC106965<br>9 AC106966<br>6 AX347067  | ALIGNMENTS 1470 bp DNA Patent W00138550. 4536200                               | Ascomycota; mitosporic 70) ing the expre  |
| %<br>Query<br>e Match Lengt | 0 100.0 147<br>5 34.4 268<br>3 10.4 218<br>2 9.3 4256 | 2 6.2 8682<br>8 5.9 10499<br>8 5.9 16954 | 5.0 1486<br>5.0 16488<br>5.0 16488<br>4.8 19292               | 4 4 7 25330<br>2 4 7 25330<br>2 4 6 11548<br>8 4 6 7341<br>2 4 6 7341<br>7 4 6 6797<br>7 4 6 19034  | 4 4 5 1043 4 6 4 6 4 6 6 4 6 6 6 6 6 6 6 6 6 6 6                                      | 14444444444444444444444444444444444444  | AX154641<br>Sequence 1 from<br>AX154641<br>AX154641.1 GI:1<br>Candida albicans | Candida albicans<br>Eukaryota; Fungi;<br>Saccharomycetales<br>1 (bases 1 to 14<br>Sundstrom, P.<br>Methods for alter<br>Patent: WO 013855 |
| Result<br>No. Scor          | 1 14<br>2 5<br>3 1<br>4 136                           | 91<br>86<br>86                           | 10 73 11 12 70 70 70 70 70 70 70 70 70 70 70 70 70            | 13 67 118 69 119 67 119 67 21 21  | 221<br>222<br>224<br>224<br>237<br>666<br>6666<br>66666666666666666666666666          | 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2   | N N  | ORGANISM O REFERENCE 1 AUTHORS 5 TITLE N JOURNAL 1  |

```
ö
                                                                                                                                                                                                                                             240
                                                                                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                                                                                                                                               480
                                                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                        61 GATAAATGTTGATTTTGTGTAATTCAATCAACTAAGCACGTTTGACAGTTAAAAAGTACG 120
                                                                                                                                                                                                     180
                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                                                                                                                                                                                        480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           780
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               780
                                                                                                                                                            61 GATAAATGTTGATTTTGTGTAATTCAATCAACTAAGCACGTTTGACAGTTAAAAAGTACG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ITCATAATTAGCCAIGTIGCACACCCTAATTCGAACATTAACTGTATCCATATTTTTCTT 720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              901 TACTATAAGCTCAAACAAATTATCTTTCAAAAATGTTATAATTAACAAGTCATCATAAT 960
                                                                                                                                          9
                                                                                                                                 TTGTTGTTGTCTCCTCGTCTCTAATTTCTGTTGACGAGGATTAATAACAGAAATACAG
                                                                                                                                                                                                                                                                                   CATAATAAAAACTCTTTGAAACATACGATATGTTATTCTTTTCATAACTGGAATATTTTT
                                                                                                                                                                                                                                                                                              GCTTTTTTTTTAACATTGAACAATTGAAAAAAAAAGGAAATGAAAAGGTAAGAGTTGCC
                                                                                                                                                                                                                                                                                                                                    GTCCTTCTTTGTTTTTTTCTAACAAAATGTTCCAGAATTTTTTAAAAAATTTTGAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GAAACCCTCCAAAAAAAAATTTTGGACCTTACACGCACATAAATTGCGGATAAACTTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CTCTCACAGTGAACTGAATTATCCATCTGAATTATCAGTCCACTAATTCCCATCAATAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TAGATTAGTGTTCTCTTCAGTACAATTACTACCATTATGCAATGCTAGCTTATTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AACACATAACACTTTGAGTATGATAATATCAACTATTGACTTGTTTTGAAAGTAAAGAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAAATTTTTTTCTAACTCGACTAATGCACTTTACATCAACTGGATGTTATTTGCATCTAC
                                                                                                                                                                                                                                                                                                                                                                                                       CAAAAGTTATTAGCGATAACCTGCGTAAGGTGTCAACAAAATATATTTGCACGTTAGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                             CTATAGAAAATATACAAACTAAATCCTTAAGGAATTTCCTCTATATAATAGGAAATCC
                                                                                                   ;
0
                                                                                Length 1470;
                                                                                                    Indels
                                                                                .
9
                                                                                                    ;
                                                                                0; DB 6;
2e-225;
                 1. 1470
/organism="Candida albicans"
/db_xref="taxon:5476"
a 260 c 186 g 515 t
                                                                              Score 1470;
Pred. No. 2e-;
Mismatches
, Paula (US)
Location/Qualifiers
                                                                                                  ;
0
                                                                               100.0%;
100.0%;
                                                                              Query Match 100.
Best Local Similarity 100.
Matches 1470; Conservative
 Sundstrom,
                                                   ď
                                                 BASE COUNT
                                                                                                                                                                                                    121
                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  199
                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                                                                                                                                             481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  781
                                                                                                                                                                                                                                                                                                                                                                                                       421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          721
         FEATURES
                                                            ORIGIN
                                                                                                                                       Dp
                                                                                                                                                            à
                                                                                                                                                                               q
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
CAUG4206 2682 bp DNA linear PLN 17-DEC-1998 Candida albicans hyphal wall protein 1 (HWP1) gene, complete cds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Staab,J.F. and Sundstrom,P. Genetic organization and sequence analysis of the hypha-specific cell wall protein gene HWPl of Candida albicans Yeast 14 (7), 681-686 (1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 (bases 1 to 2682)
Staab.J.F. and Sundstrom,P.
Direct Submission
Submitted (18-JUL-1996) Medical Microbiology & Immunology, Ohio State University, 333 West 10th Avenue, Columbus, OH 43210, USA 4. (bases 1 to 2682)
                                                                                             1080
                                                                                                                         1080
                                                                                                                                                      1140
                                                                                                                                                                                                           1200
                                                                                                                                                                                                                                     1200
                                                                                                                                                                                                                                                                  1260
                                                                                                                                                                                                                                                                                            1260
                                                                                                                                                                                                                                                                                                                        1320
                                                                                                                                                                                                                                                                                                                                                                             1380
                                                                                                                                                                                                                                                                                                                                                                                                       1380
                                                                                                                                                                                                                                                                                                                                                                                                                                   1440
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Saccharomycetales; mitosporic Saccharomycetales; Candida.
1 (bases 1 to 2682)
Staab,J.F., Ferrer,C.A. and Sundstrom,P.
Developmental expression of a tandemly repeated, proline-and glutamine-rich amino acid motif on hyphal surfaces on Candida albicans
                                                                                                                                                                                                          1021 ATTGTAAAAAGGGAGAGTTTTGGTAGGCTCATAATCGCTTATAATGTTACCTCTAAAGTAA
                                                                                                            1021 ATTGTAAAAAGGGAGAGTTTTGGTAGGCTCATAATCGCTTATAATGTACCTCTAAAGTAA
                                                                                                                                                                                                                                                                                                                         1381 ATTTTCAATTTCCATTCAACTTGTTTTTCTCAACAATATCAAACACAACACAGGAATCTCCTA
                                         961 TCTTTGGATCCAAAAACAAGGAATTCGGAAATTCTGACGATAAATGTCGACTCACAATTC
                                                                                                                                                      CACTITITGTIGCGACTITAATACCGTTTTTGCAACTICTCTTTGTATCACCTGTATCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              J. Biol. Chem. 271 (11), 6298-6305 (1996) 96198091
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1441 TAGTCACTCGCTTTTAGTTTCGTCAATATG 1470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1441 TAGTCACTCGCTTTTAGTTTCGTCAATATG 1470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                U64206.1 GI:4028879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Candida albicans.
Candida albicans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8626424
                                                                                                                                                                                                                                                                  1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 2
CAU64206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LOCUS
                                         ò
                                                                     qq
                                                                                             δ
                                                                                                                         g
                                                                                                                                                                              pp
                                                                                                                                                                                                           QΥ
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                         ōλ
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                      ò
                                                                                                                                                                                                                                                                  Qγ
                                                                                                                                                                                                                                                                                                                                                                               ò
```

1446 ACTCGCTTTTAGTTTCGTCAATATG 1470

ŏλ

```
SSTTPRTSASTTPESSYPATTPNTSVPTTSSESTTPATSPESSYPVTSGSSILATTSE
SSSAPATTPNTSVPTTTETKSSSTPLTTTTEHDTTVVTTSCSNSVCTESEVTTGVI
VITSKDITTTTTTTTCPLTETKSSSTPLTTTTTEHDTTVVTTSCSESSCTES
BVTTGVVVVTSEETVYTTRCPLTENTPGTDSTPEASIPPMETIPAGSESSESTES
AVPKSDVPATESAPVPTTFCPLTENTPGTDSTPEASIPPMETIPAGSESSAPAGETSP
AVPKSDVPATESAPAPEMTPAGSQPSIPAGETSPAVPKSDVPATESAPAPEMTPAGTS
SCAITIPESSAVVSTTEGAIPTLESVPLMQPSANXSSVAPISTFEGAGNNMRLTFGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                /translation="MRLSTAQLIAIAYMLSIGATVPOVDGOGETEEALIQKRSYDYY
QEPCDDYPQQPQEPCDYPQOQQEEPCDYPQQPQE
PCDYPQQPQEPCDNPPQPDVPCDNPPQPDVPCDNPPQPDIPCDNPPQP
DQPDDNPPIPNIPTDMIPNIPTDMIPDIPEKPTTPATTPATTTSESSSSSSSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GTTAGTTAGCCCAGCTGTTTTTTTTTTGCCTTATTTTTTAGACTACATTTGTTTCACTT 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAAAAGGGAGAGTTTTGGTAGGCTCATAATCGCTTATAATGTACCTCTAAAAGTAATCTAA 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AACAAACACAACCTTTCTAAAACCTATAATAATAACCCTAATGGCTCACAAACGGGGTAA 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITGITGCGACTITAATACCGITTITGCAACTICTCTTTGTATCACCTGTATCCGCCTTT 1265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TTTAACATAGCAACTCTTGTAAAGTCCCTTTCTTTTCCCACTATTTTATCATTCTTGAAA 1325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TATGTAATCAGAATAGTTTTTCAAAAACTATAAATAACGGTCAAAATAACCGGCTATTTT 1385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1386 CAATTTCCATTCAACTTGTTTTCTCAACAATATCAAACACAACAGGAATCTCCTATAGTC 1445
            Submitted (17-DEC-1998) Medical Microbiology & Immunology, Ohio State University, 333 West 10th Avenue, Columbus, OH 43210, USA Sequence update by submitter On Dec 17, 1998 this sequence version replaced gi:1915978.

Location/Qualifiers
1. .2682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   966 GGATCCAAAAACAAGGAATTCGGAAATTCTGACGATAAATGTCGACTCACAATTCATTGT 1025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          421 CAATTTCCATTCAACTTGTTTTCTCAACAATATCAAAACACAAAACTCCCTATAGTC 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241 ITIGITGCGCCTTTAATACCGTTTTTGCAACTTCTCTTTGTATCACCTGTATCGCCCTTT 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GTTAGTTAGCCCAGCTGTTTTTTTTTGCCTTATTTTTATGACTACATTTTGTTTCACTT 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            361 TATGTAATCAGAATAGTTTTTCAAAAACTATAAATAACGGTCAAAATAACCGGCTATTTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34.4%; Score 505; DB 8; Length 2682; 100.0%; Pred. No. 1.8e-71; Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                     /product-"hyphal wall protein 1"
/protein_id="AAC96368.1"
/db_xref="G1:1915979"
                                                                                                                                                                                                                                                                                                        /gene="HWP1"
/note="hyphal surface protein"
                                                                                                                                          /organism="Candida albicans"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        875 t
                                                                                                                                                                            /db_xref="taxon:5476"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      370 g
                                                                                                                                                                                                                                                                                                                                                                      /transl_table=12
                                                                                                                                                           /strain="SC5314"
                                                                                                                                                                                                                                                                                                                                                 /codon_start=1
                                                                                                                                                                                                                                /gene="HWP1"
503. 2682
                                                                                                                                                                                                                     /gene="HWP1"
445. .2682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AIIGIAAFLI"
678 c
                                                                                                                                                                                                  .2682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 505; Conservative
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
  TITLE
JOURNAL
                                                                                                                                                                                                                                               mRNA
                                                                                                                                                                                                           gene
                                                                                                                                                                                                                                                                                         CDS
                                                                                                   FEATURES
                                                              REMARK
                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
```

```
/u_a_au_i - u_i : z'i / J3A
/u_au_i - u_i : z'i / J3A
/u_au_i - u_i : z'i / J3A
/u_au_i - u_i / J3A
/u
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                           PLN 24-JUL-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sharkey, L.L., Saporito-Irwin, S.M. and Fonzi, W.A.
Direct Submission
Submitsed (01-MNY-1997) Microbiology and Immunology, Georgetown University, 3900 Reservoir Road NW, Washington, DC 20007-2197, USA on Jul 24, 1997 this sequence version replaced gi:2209195.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note="differentially expressed in relation to the extent
                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Candida. [ (bases 1 to 2188) Saccharomycetales; Candida. Sharkey, L. L., Saporito-Irwin, S.M. and Fonzi, W.A. ECE2, a differentially expressed gene of Candida albicans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 ICTIGAAATAIGIAATCAGAATAGTTITICAAAAACTATAAATAAGGGTCAAAATAACCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 2188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10.4%; Score 153; DB 8; I
100.0%; Pred. No. 2.8e-15;
tive 0; Mismatches 0;
                                                                                                                                                                                                                              Candida albicans ECE2 gene, complete cds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism="Candida albicans"
/strain="SC5314"
/db_xref="taxon:5476"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        670 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CTATAGTCACTCGCTTTTAGTTTCGTCAATATG 1470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CTATAGTCACTCGCTTTTAGTTTCGTCAATATG 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /transl_table=12
/protein_id="AAB64014.1"
/db_xref="G1:2275336"
                                                                                                                                                                                                      2188 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 of cell elongation"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        306 g
481 ACTCGCTTTTAGTTTCGTCAATATG 505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /chromosome="4"
151. .2055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon_start=1
                                                                                                                                                                                                                                                                                                  AF001978.1 GI:2275335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /gene="ECE2"
151. .2055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          151. .2055
/gene="ECE2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AIIGIAAFLI"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ECE2, a differential!
Unpublished
2 (bases 1 to 2188)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               587 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .2188
                                                                                                                                                                                                                                                                                                                                                                            Candida albicans
                                                                                                                                                                                                                                                                                                                                                                                                            Candida albicans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                      AF001978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                      RESULT 3
AF001978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                            KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMMENT
                                                                                                                                                                                                                                                                                                            VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                            SOURCE
                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                   g
```

RESULT 4

TITLE

SOURCE

```
misc_feature
                                                                                                                   misc_feature
                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                              aeue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                   PLN 05-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (05-NOV-1998) On behalf of the pilot sequencing project on the Candida albicans strain 1161 genome. Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CBIO 1SA B-mail:
Darrell@sanger.ac.uk Cosmids supplied by Prof. Duncan Shaw, [3]
Department of Molecular and Cell Biology, The Institute of Medical Science, University of Aberdeen, Foresterhill, Aberdeen, AB25 2ZD,
                                                                                                                                                       ARS; Canikl; CDP-alcohol phosphatidyltransferase; chitin synthase; chs1; cytochrome P450; DNA polymerase delta; histidine kinase; isocitrate dehydrogenase pseudogene; LTR; pol3; Rcc1; regulator of chromosome condensation; rehydrin; RPS1-like region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /translation="MOLFIPPNKLQNLKLYKYSSEDHSIISKYILKKWWNFFYQIFPL
SMAPNVVTLLGLFFIIGNLMTVFYYDDYLNETQPTWCYFFYAFGLFWYQTFDGCDGCH
ARRTGQSGPLGELFDHSIDAINTTLGTFVFASVLKMGYGGLLLLSQFASVCNFYTSTW
EEYHTHTLFLSKFSGPVEGILMICIVYIITGIFGPDIWTIDLFELNLTSLGYGYYKVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Complement(41. 785)

Complement(41. 785)

Complement(41. 785)

Complement(51. 785)

Anote="Ca35A5.01c," partial orf, len: > 261 aa, most

Similar to EPT1_YEAST ethanolaminephosphotransferase (EC 2.7.8.1) (391 aa), fasta scores opt: 984, E(): 0, (56.38 identity in 263 aa overlap), contains PS00379 CDP-alcohol phosphatidyltransferases signature"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Funding: sequencing funded by Beowulf Genomics Ltd. CDS are numbered using the following system eg CAC2CCI.OIC. CA (C. albicans), 20cI (cosmid name),
01 (first CDS), c (complementary strand).
The more significant matches with motifs in the PROSITE/PFAM database are also included but some of these may be fortuitous. IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence soverlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid Ca35AS is likely to map to region R, chromosome 7.
                                                                                                                                                                                                                                                                                                                                 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.

(bases 1 to 42565)

Tait, E., Simon, M.C., King, S., Brown, A.J., Gow, N.A. and Shaw, D.J. A Candida albicans genome project: cosmid contigs, physical mapping, and gene isolation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /product="putative alcohol phosphatidyl transferase"
                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism="Candida albicans"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /protein_id="CAA21944.1"
/db_xref="GI:3850144"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 (bases 1 to 42565)
Barrell, B.G. and Rajandream, M.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone="cosmid Ca35A5"
complement(1. .785)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /strain="1161"
/db_xref="taxon:5476"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /map="chromosome 7"
                                                                     C.albicans cosmid Ca35A5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /transl_table=12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Oliver, K. and Harris, D.
                                                                                                                                  GI:3850143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Article No. FG970983
2 (bases 1 to 42565)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1. .42565
                                                                                                                                                                                                                                                                           Candida albicans.
Candida albicans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
                                                                                                                                  AL033396.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Notes:
                                          LOCUS
                                                                                           ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE
PUBMED
               CAC35A5/c
                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMARK
```

COMMENT

FEATURES

```
MYEIERIKKANSEYROSIDKIKEILIYNDOKIDSONBAFSSLAKSALKQVDKNYQLKI
19043. 19210
/note="Blastw match (score: 669, strand:(+1), range:
8-175) to AF069450 C. albicans retrotransposon long
reminal repeat zeta"
19204. 19277
/note="Blastw hit (score: 325, strand: (+1), range:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SESSKSVSGNRLRIAKKLKFSKNQSSRKPDLFLDTSSTISEDSSSFRKVIIGNSLNEM
TKPRKSSMSSPMSTTFHSLFHRSHHNGSNLQRDTNQVATGTTPLSGKFDDFSKASKTT
LCLSSNSSNSINSIISNPELAQIYNFTNPNISIEDRETNLDHTNSSFLDIHKKMLVPADSF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IQNKLNKYHQTEVGLGIYESELDHDNDNKIYSNLYHYLKPLFTPSFSISDSGOKSKMR
PITSANVERETNRFWESFCLHOPNHERSFRSKYRKSVSSIGENGKNVEDEPPYLQLSNLFP
KLMSLLSHNLQTTBESVSYLGALILNAMKYYNAYVRPTLSTFQPLQTYLNEFTRSH
NGSKIIRIDDLLLVSFRKVFITEQGIGSGERETSQFLGNAESEDLTGNGLLTSTLAKL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /protein_id="CAA21946.1"
/db_xref="G1:3850146"
/translation="MNQSNNKSNTERRNSIEEEDALHYHYHEYSVSLVQHEKNKQEIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /translation="MSEVPNQESPSSSILKQSNASSSTKVHGAPSSSQIANVDFNTLR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KLNSNTSFSSNLTNSTTNTSNALSRLFTRNKSLSNISVHPSSSDDDSSPKTLRESTSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /gene="Ca35A5.03"
18031. .18456
/gene="Ca35A5.03"
/note="Ca35A5.03, unknown, questionable orf, len: 141 aa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /jote="Ca35A5.04, Chs1 gene, len: 1026 aa, CHS1_CANAL chitin synthase 1 (EC 2.4.1.16), CHS1_CANAL (776 aa) (99.9% identity in 776 aa overlap) but open reading frame here extends extends 250 aa upstream of CHS1_CANAL start
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="BLASTN match (score: 1704 strand: (+1), range: 49-459) to CAAURESE X65035 C.albicans DNA of an autonomously replicating sequence (ARS)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        168-24) to AF069450 C. albicans retrotransposon long terminal repeat zeta" 20359. .20769
                                                                            /note="PS00379 CDP-alcohol phosphatidyltransferases
                                                                                                                                                                                                                         /note="region showing similarity with isocitrate dehydrogenase"
                                                                                                                                                                                                                                                                                                                                                          /gene="ca35A5.02"
/mote="Ca35A5.02, unknown, len: 512 aa, possible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TSIIYTIIGLTSLYFNIASAMFNVSKHYKKSSTNNSSSGDKDGS"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /product="hypothetical protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /product="questionable orf"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /transl_table=12
/product="chitin synthase"
/protein_id="CAA21947.1"
                                                                                                                                       4307. .11124
/note="RPS1-like region"
complement(7606. .8085)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /protein_id="CAA21945.1<sup>†</sup>
/db_xref="GI:3850145"
                          complement(423. .491)
/qene="Ca35A5.01c"
                                                                                                                                                                                                                                                                              13175. . 14713
/gene="Ca35A5.02"
13175. .14713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /transl_table=12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /transl_table=12
                                                                                                                                                                                                                                                                                                                                                                                                                                                     /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .18456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20808. .23888
/gene="Chs1"
                                                                                                                                                                                                                                                                                                                                                                                                                  glycoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /gene="Chs1"
                                                                                                         signature"
4307. .111
```

```
January 3, 2003, 19:47:58 ; Search time 15.5 Seconds (without alignments) 94.913 Million cell updates/sec
                                                                                                                                                                                                                                  US-09-725-010-3
281
1 CINCFIQITPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                           262574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Issued_patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/pacycTUS_COMB.pep:*
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                           262574 seqs, 29422922 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                            OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                            Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                      BLOSUM62
                                                                                                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                       Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Database :
                                                                                                                                                                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                         Sequence:
                                                                                                                                        Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

|           | Description    | Sequence 2, Appli | 9               | 7                | Sequence 2, Appli | 11,              | 11,               | 11,              | 11,              | 11,               | 48,               | 48,               | 8,              | 147                | П                | 19,              | 8, 1            | 'n               | Sequence 4, Appli | 11,               | 11,              | Sequence 37, Appl | 37,              | 37,              | 36,               | 36,              | 36,     | Sequence 13, Appl |
|-----------|----------------|-------------------|-----------------|------------------|-------------------|------------------|-------------------|------------------|------------------|-------------------|-------------------|-------------------|-----------------|--------------------|------------------|------------------|-----------------|------------------|-------------------|-------------------|------------------|-------------------|------------------|------------------|-------------------|------------------|---------|-------------------|
| SUMMARIES | ID             | US-09-254-325-2   | US-09-079-415-6 | US-08-750-458A-2 | US-08-927-394-2   | US-08-204-740-11 | US-09-081-167A-11 | US-09-081-395-11 | US-09-416-833-11 | PCT-US95-02521-11 | US-08-455-543A-48 | US-08-223-305C-48 | US-09-268-163-8 | PCT-US95-06266-141 | US-08-460-309-19 | US-08-125-077-19 | US-08-852-153-8 | US-09-103-429A-3 | US-09-103-429A-4  | US-08-390-874C-11 | US-09-265-772-11 | US-08-162-081B-37 | US-08-780-872-37 | US-09-085-957-37 | US-08-162-081B-36 | US-08-780-872-36 | -09-082 | US-08-269-441A-13 |
|           | DB             |                   | ٣               | m                | 7                 | Н                | ٣                 | m                | 4                | S                 | Н                 | 7                 | 4               | S                  | 7                | 7                | 7               | 4                | 4                 | ٣                 | 4                | ~1                | ~                | 4                | 7                 | 7                | 4       | -                 |
|           | Length DB      | 853               | 866             | 866              | 449               | 108              | 108               | 108              | 108              | 108               | 2237              | 2237              | 2237            | 128                | 314              | 314              | 664             | 786              | 802               | 1068              | 1068             | 1069              | 1069             | 1069             | 1080              | 1080             |         | 63                |
| ф         | Query<br>Match | 98.6              | 98.6            | 98.6             | 66.99             | 18.9             |                   | 18.9             | 18.9             | 18.9              | 18.7              | 18.7              | 18.7            | 18.5               | 17.4             | 17.4             | 17.4            | 17.4             | 17.4              | 17.4              | 17.4             | 17.4              | 17.4             | 17.4             | 17.4              | 17.4             | 17.4    | 17.3              |
|           | Score          | 277               | 277             | 277              | 188               | 53               | 53                | 53               | 53               | 53                | 52.5              | 52.5              | 52.5            | 52                 | 49               | 49               | 49              | 49               | 49                | 49                | 49               | 49                | 49               | 49               | 49                | 49               | 49      | 48.5              |
|           | Result<br>No.  | 1                 | 7               | m                | 4                 | Ŋ                | φ                 | 7                | œ                | σ                 | 10                | 11                | 12              | 13                 | 14               | 15               | 16              | 17               | 18                | 19                | 20               | 21                | 22               | 23               | 24                | 25               | 26      | 27                |

| ce 2, Applice 3, Applice 3, Applice 3, Applice 3, Applice 3, Applice 2, Applice 2, Applice 2, Applice 114, Applice 114, Applice 114, Applice 114, Applice 2, Applice 3, Applice   |  | Gaps 0;   |
|---|--|---|
| sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sedneuce<br>Sed | OLDS (EPO)   | ED BY<br>ED BY<br>ngth 853;<br>Indels 0;<br>KKRNR 50<br>     <br>KKRNR 701  |
| 314 1 US-08-269-441A-2 314 4 US-09-008-465-3 314 4 US-09-276-851-2 314 4 US-09-276-851-2 314 4 US-09-276-851-2 314 4 US-09-38-895-390-3 785 4 US-08-895-590-3 785 4 US-08-895-590-2 2516 4 US-08-895-590-2 2516 4 US-09-539-879A-2 2516 4 US-09-579-181-1 256 2 US-08-484-905-114 256 3 US-08-481-9858-114 256 4 US-09-579-181-1 256 3 US-08-481-9858-114 256 4 US-09-579-181-1 257 4 US-09-579-181-1 258 3 US-08-481-9858-114 259 4 US-09-579-181-1 251 4 US-09-579-181-1 252 4 US-08-370-476-114 253 4 US-08-370-476-114 254 4 US-09-579-181-1  | MES IN KOJI M  | /note= "DNA BINDING SITE" /note= "TRUNCATED AREA WHICH STILL ACTIVE BUT NOT REPRESS L-GLUTAM" 98.6%; Score 277; DB 3; Le 98.0%; Pred. No. 1.3e-30; ive 1; Mismatches 0; DGGPLCNACGLFLKLHGVVRPLSLKTDVI :                            'EGGPLCNACGLFLKLHGVVRPLSLKTDVI 'EGGPLCNACGLFLKLHGVVRPLSLKTDVI 'EGGPLCNACGLFLKLHGVVRPLSLKTDVI 'EGGPLCNACGLFLKLHGVVRPLSLKTDVI 'EGGPLCNACGLFLKLHGVVRPLSLKTDVI 'EGGPLCNACGLFLKLHGVVRPLSLKTDVI 'EGGPLCNACGLFLKLHGVVRPLSLKTDVI |
| 48.5 17.3 33.0 48.5 17.3 48.5 17.3 48.5 17.3 33.3 48.5 17.3 48.5 17.3 24.6 48.5 17.3 24.6 48.5 17.3 24.0 48.5 17.3 44.0 48.5 17.1 43.5 44.5 44.5 44.5 44.5 44.5 44.5 44.5   | RESULT 1 US-09-254-325-2 Sequence 2, Application US/09254325 Patent No. 6090607 TITLE OF INVENTION: ENHANCED EXPRESS TITLE OF INVENTION: PROTECUTIC ENZINUMBER OF SEQUENCES: 17 COMPUTER READABLE PORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC Compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PATENTIN Relase #1.0, NC CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/254,325 INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS: LENGTH: 853 amino acids TYPE: amino acid STRANDEDNESS: single TYPE: amino acid STRANDEDNESS: single TYPE: MOMECULE TYPE: protein FEATURE: NAME/KEY: Binding-site | COCATION: 652-676   |
| ·   | й м · · · · · · · · · · · · · · · · · ·  | O V V V V V V V V V V V V V V V V V V V   |

us-09-725-010-3.rai

ö

Gaps

ö

Indels

```
TITLE OF INVENTION: GATA-6 TRANSCRIPTION FACTOR: COMPOSITIONS AND METHODS NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS: ADDRESSEE: Wolf, Greenfield, & Sacks, P.C.
STREET: 600 Atlantic Avenue
                                                                                                                                                                                                                                                                                                                                                             Score 277; DB 3; Length 866; Pred. No. 1.4e-30;
                                                                                                                                                                                                                                                                                                                                       1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OPERATING SYSTEM: PC-DOS/MS-DOS
SOPTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       66.9%; Score 188; DB 2;
illarity 62.0%; Pred. No. 2.1e-18;
Conservative 9; Mismatches 10;
                                                                                                                                                                                                                                                                                              1; Mismatches
  4129.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S1237/7005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFCATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/025,574
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; Sequence 2, Application US/08927394
; Patent No. 5990092
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36,637
REFERENCE/DOCKET NUMBER: 41
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 867 0123
TELEFAX: 212 867 0298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: Plumer, Elizabeth R. REGISTRATION NUMBER: 36,63 REFERENCE/DOCKET NUMBER: S
                                                                                                                                                                                                                                                     98.6%;
98.0%;
                                                              TELEFAX: 212 867 0298
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                           LENGTH: 866 amino acids TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : 449 amino acids amino acid
                                                                                                                                                                                                                                                Query Match
Best Local Similarity 98.0°
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                  ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-750-458A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        , MOLECULE TYPE: protein US-08-927-394-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best_Local Similarity
Matches 31; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-927-394-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENERAL INFORMATION:
APPLICANT: Christensen, Tove
APPLICANT: Hynes, Michael J.
TITLE OF INVENTION: A Fungus Wherein The Area Gene Has Been Modified And An Area
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           E: No. 60251850 No. 6025185disk of No. 6025185th America, Inc. 405 Lexington Avenue, Suite 6400
                                                                                                                           No. 60134520 No. 6013452disk of No. 6013452th America, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                   APPLICANT: Lehmbeck, Jan
TITLE OF INVENTION: A Fungus Wherein The area, pepC and/or
TITLE OF INVENTION: pepE Genes Have Been Inactivated
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREEF: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
CUUTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: DE PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   664 CTNCFTQTTPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 866;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                              COMPOTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,415
FILING DATE: 14 -MAY-1998
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 277; DB 3;
Pred. No. 1.4e-30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4657.204-US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/08/750,458A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 2, Application US/08750458A Patent No. 6025185
                                                                                                                                               405 Lexington Avenue
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3-December-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY AGENT INFORMATION:
NAME: ROZEK, Carol E.
REGISTRATION NUMBER: 36,993
                                                                                                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: ROZEK, CATOL
REGISTRATION NUMBER: 36,993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 46:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEPAX: 212-878-9655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR SEQ ID NO: 6: SEQUENCE CHARACTERISTICS: LENGTH: 866 amino acids TYPE: amino acids
  Christensen, Tove
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98.6%;
98.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TOPOLOGY: linear; MOLECULE TYPE: protein US-09-079-415-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
                                                                                                                                                                    New York
                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                             10174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE:
                                                                                                                             ADDRESSEE:
                                                                                                                                                                                        Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 3
US-08-750-458A-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    49;
                                                                                                                                                                                                           COUNTRY:
                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
Gaps
.;
0
                                  1 CINCFIQITPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
```

Length 449;

```
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE:
STREET: 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-081-395-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pp
                                                            GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Razarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Mazo, Ilya
APPLICANT: Molison, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Gudkov, Andrei
APPLICANT: Gudkov, Andrei
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Mazo, Ilya
APPLICANT: Moligian, Igor B
APPLICANT: Moligian, Igor B
APPLICANT: Moligian Interest Invention: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18.9%; Score 53; DB 1; Length 108; 43.5%; Pred. No. 3.8; ative 7; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 S. Wacker Drive, 32nd Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIUM TYPE: Floppy disk
COMEUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/204,740
FILING DATE: 04-MAR.1994
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                 93,354-C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION:
NAME: No. 5753432na, Kevin E
REGISCHAFTON NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-1
TELECOMMUNICATION INFORMATION:
TELECAX: 312-715-100
TELEEXX: 312-715-100
TELEEXX: 910-221-5317
INFORMATION FOR SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-081-167A-11
Sequence 11, Application US/09081167A
Patent No. 6083745
GENERAL INFORMATION:
                       ; Sequence 11, Application US/08204740
; Patent No. 5753432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28 FLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14 FLEANEVPRPVTLRTNTLKTRRR 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; MOLECULE TYPE: peptide US-08-204-740-11
                                                                                                                                                                                                                                                                      CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 10; Conserv
                                                                                                                                                                                                                                                                                                                                                           STATE: Illinois
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: Chicago
STATE: Illinois
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                   90909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZIP: 60606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
US-08-204-740-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
```

```
APPLICANT: Gudkov, Andrei
APPLICANT: Gudkov, Alexander
APPLICANT: Kazarov, Alexander
APPLICANT: Mazo, Ilya
APPLICANT: Roninson, Ilya
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
CORRESPONDENCES 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
18.9%; Score 53; DB 3; Length 108;
Best Local Similarity 43.5%; Pred. No. 3.8;
Matches 10; Conservative 7; Mismatches 6; Indels
                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OOPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,167A
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 6083745nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEFAX: 312-913-0002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             E: McDonnell Boehnen Hulbert & Berghoff
300 S. Wacker Drive, 32nd Floor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 18-MAY-1998
CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
NAME: No. 6083746nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-KK
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 11, Application US/09081395
Patent No. 6083746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28 FLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR SEQ ID NO: 11:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 312-913-0002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; MOLECULE TYPE: peptide US-09-081-167A-11
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: Chicago
STATE: Illinois
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GENERAL INFORMATION:
```

```
ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
US-08-455-543A-48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STREET:
                                                      APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Razarov, Alexander
APPLICANT: Razarov, Alexander
APPLICANT: Mazo, 11ya
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods for Identifying Genetic
TITLE OF INVENTION: Suppressor Elements and Genes Associated with Malignant
TITLE OF INVENTION: Growth in Cancer Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
                                                                                                                                                      ő
                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0
                                                                                                                                                      0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18.9%; Score 53; DB 4; Length 108; illarity 43.5%; Pred. No. 3.8; Conservative 7; Mismatches 6; Indels
                                                                                                               Length 108;
                                                                                                                                                      6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READBLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,833
                                                                                                               Query Match 18.9%; Score 53; DB 3;
Best Local Similarity 43.5%; Pred. No. 3.8;
                                                                                                                                                    7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: Allgretti & Witcoff, Ltd.
STREET: 10 S. Wacker Drive, Suite 3000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 6197521nan, Kevin E
REGISTRATION NUMBER: 35,303
REGISTRATION NUMBER: 93,354-C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CLASSIFICATION:
PARIOR APPLICATION DATA:
APPLICATION NUMBER: 08/204,740
FILING DATE:
                                                                                                                                                                                                                                                                                                                                 Sequence 11, Application US/09416833 Patent No. 6197521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 93,3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1200
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 11:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28 FLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14 FLEANEVPRPVTLRINTLKTRRR 36
                                                                                                                                                                                             28 FLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                   14 FLEANEVPRPVTLRTNTLKTRRR 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        : 108 amino acids
amino acid
LENGTH: 108 amino acids
                                                                                                                                                      10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS
                          MOLECULE TYPE: peptide US-09-081-395-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; MOLECULE TYPE: peptide US-09-416-833-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDALL
STREET: 10 C
CITY: Chicago
                    amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best_Local Similarity
Matches 10; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: US
ZIP: 60606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY:
                                                                                                                                                                                                                                                                                        RESULT 8
US-09-416-833-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                      Matches
                                                                                                                                                                                                                                 qq
                                                                                                                                                                                             ò
```

```
Methods for Identifying Genetic
Suppressor Elements and Genes Associated with Malignant
Growth in Cancer Cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Anniel
APPLICANT: McCue, Anniel
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
CORRESPONDENCES: 57
CORRESPONDENCES: 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18.9%; Score 53; DB 5; Length 108; 43.5%; Pred. No. 3.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6; Indels
                                                                                                                                                                                                                                                              OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Brown, Martin, Haller & McClain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,543A
FILING DATE: MAY 31, 1995
FILING DATE: MAY 31, 1995
APPLICATION NUMBER: 08/23,305
FILING DATE: APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: APTI1 10, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: APTI1 10, 1992
PRIOR APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-ANG-1991
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-ANG-1991
PRIOR APPLICATION DATA:
PCT-US95-02521-11; Sequence 11, Application PC/TUS9502521; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 48, Application US/08455543A

Patent No. 5792846

GENERAL INFORMATION:

APPLICANT: Harpold, Michael

APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark
                                                                                    TITLE OF INVENTION: Methods for TITLE OF INVENTION: Suppressor I TITLE OF INVENTION: Growth in C2 NUMBER OF SEQUENCES: 13 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28 FLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
INFORMATION FOR SEQ ID NO: 11:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1660 Union Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                              : 108 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 43.5
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; TOPOLOGY: linear; MOLECULE TYPE: peptide PCT-US95-02521-11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: San Diego
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
```

APPLICATION NUMBER: US 07/620,250

RESUL,T 9

Mon Jan

```
FILING DATE: 15-AUG-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-223-305C-48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH: 2237
TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-268-163-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-268-163-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 18.7%; Score 52.5; DB 1; Length 2237; Best Local Similarity 37.3%; Pred. No. 1.7e+02; Matches 19; Conservative 3; Mismatches 22; Indels 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5 FTQTTPLWRRNPDGQPLCNACGLFLKLH----GVVRPLS---LKTDVIKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Brown, Martin, Haller & McClain
                                                                                                                                                                   PRIOR AFFILIATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 33,779
FELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 2237 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: US/08/223,305C FILING DATE: API14, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: API11, 1992
PRIOR APPLICATION DATA:
APPLICATION DATA: US 07/745,206
                                        APPLICATION NUMBER: US 07/482,384 FILING DATE: 20-FEB-1990 PRIOR APPLICATION DATA:
                                                                                                             US 07/603,751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 48, Application US/08223305C Patent No. 5851824 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER: IBM Compatible OPERATING SYSTEM: DOS
                                                                                                        APPLICATION NUMBER: US 07
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: USA ZIP: 92101-2926 COMPUTER READABLE FORM: MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE TYPE: protein FRAGMENT TYPE: internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  single
                    PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
STRANDEDNESS: Sin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-455-543A-48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-08-223-305C-48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
```

```
APPLICANT: Lipscombe, Diane
APPLICANT: Schorge, Stephanie
TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
FILE REFERENCE: B1055/7000
CURRENT APPLICATION NUMBER: US/09/268,163B
CURRENT FILING DATE: 1999-03-12
EARLIER FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FASTSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2170 FPNTTPCCRETPSASPWPLALELALTLTWGSVWTVRPLSTPCLRTRSLSRR 2220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 18.7%; Score 52.5; DB 2; Length 2237; Best Local Similarity 37.3%; Pred. No. 1.7e+02; Matches 19; Conservative 3; Mismatches 22; Indels 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db 2170 FPNTTPCCRETPSASPWPLALELALTLTWGSVWTVRPLSTPCLRTRSLSRR 2220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18.7%; Score 52.5; DB 4; Length 2237; 37.3%; Pred. No. 1.7e+02;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5 FTQTTPLWRRNPDGQPLCNACGLFLKLH----GVVRPLS---LKTDVIKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 FTQTTPLWRRNPDGQPLCNACGLFLKLH----GVVRPLS---LKTDVIKKR 48
                                                                    PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/482,384
FILING DATE: 20.FEB-1990
PRIOR APPLICATION NUMBER: US 07/603,751
FILING DATE: 04.APR-1989
FILING DATE: 04.APR-1989
FILING DATE: 04.APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04.APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04.APR-1988
ATTORNEY.AGENT INFORMATION:
NAME: Sendman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELEPHONE: (619)238-0999
TELEPHONE: (619)238-0662
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENTH: 2237 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3; Mismatches
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 8, Application US/09268163B
Patent No. 6353091
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein FRAGMENT TYPE: internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: sir
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
```

```
5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18.5%; Score 52; DB 5; Length 128; 31.2%; Pred. No. 6.5; tive 10; Mismatches 21; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  37 CMMPSTRQTPLYPRWQDTKGIPGSCGMSLAFSQVLK--SLNTSHIQSQ 82
                                                                                                   Detection of Viral Antigens Coded
by Reverse Reading Frames
157
                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06266
                                                                                                                                                                                                    ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4600-0202.41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE: 20-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/285,561
FILING DATE: 03-MG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,729
FILING DATE: 26-CCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,271
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/357,509
FILING DATE: 16-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/357,509
FILING DATE: 16-DEC-1994
PRIOR APPLICATION NUMBER: US 08/357,509
FILING DATE: 15-DEC-1994
APPLICATION NUMBER: US 08/357,509
FILING DATE: 15-DEC-1994
APPLICATION NUMBER: US 08/357,509
FILING DATE: 15-DEC-1994
APPLICATION NUMBER: US 08/389,886
FILING DATE: 15-FEB-1995
ATTORNEY AGENT INFORMATION:
                        Sequence 141, Application PC/TUS9506266 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                          COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: M62321 ORF1
PCT-US95-06266-141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0860
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Fabian, Gary R. REGISTRATION NUMBER: 33,875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE/DOCKET NUMBER: 46
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 31.28
Matches 15; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: protein HYPOTHETICAL: NO ANTI-CENT
                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
                                                                                                TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                            Palo Alto
                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                  94306
PCT-US95-06266-141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
US-08-460-309-19
                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                           APPLICANT
                                                                                                                                                                                                                                                                                    STATE:
```

Sequence 19, Application US/08460309 Patent No. 5837496

```
12; Indels 24; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNC---FTQTTPLWRRNPDGQPL-CNAC------GLFLKLHG 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
CORRESPONDENCE: 23
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INCORATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego.
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 17.4%; Score 49; DB 2; Length 314; Best Local Similarity 27.3%; Pred. No. 50; Matches 15; Conservative 4; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                       SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,309
FILING DATE:
                                                                                                                                                                        ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION NUMBER: US 07/472,319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION NUMBER: US 07/919,951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REGISTRATION NUMBER: 9-LA 9721
TELECOMMUNICATION INDER: P-LA 9721
                                                                                                                                                                                                                                                                                                   ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ; Sequence 19, Application US/08125077
; Patent No. 5872231
; Patent No. 5872231 5840863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 314 amino acids
                                                                                                                                                                                                                                                 California
: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TOPOLOGY: linear
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
US-08-125-077-19
                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                            CITY: S
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
```

```
COMPUTER RADABLE FORM:

COMPUTER: IN PC compatible

CLASSIFICATION UNDER: US/08/125,077

FILING DATE: 22-SEP-1994

RIOR APPLICATION UNDER: US PCT/US 94/10730

FILING DATE: 30-TAN-1990

PRIOR APPLICATION UNDER: US 07/412,319

FILING DATE: 30-TAN-1990

ATTORNEY AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REFERRATION NUMBER: 13,815

RECEDENCE/DOCKET NUMBER: P-LA 9721

FILING DATE: 31,815

RECEDENCE/DOCKET NUMBER: P-LA 9721

FILING DATE: 31,815

RECEDENCE/DOCKET NUMBER: 19,910

FILING CHARACTERISTICS: 19,910

FELERAX: (619) 535-9001

FELERAX: (619) 535-9001

FELERAX: (619) 535-9001

FELERAX: (10) 335-900

FEL
```

Search completed: January 3, 2003, 19:52:38 Job time: 16.5 secs

|   |     | *  |     |   |
|---|-----|----|-----|---|
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     | */ |     |   |
|   | (4) |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    | -35 |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
| 3 |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     |   |
|   |     |    |     | • |
|   |     |    |     |   |
|   |     |    |     | • |

```
GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

January 3, 2003, 19:46:47; Search time 25 Seconds (without alignments) 192.269 Million cell updates/sec Run on:

US-09-725-010-2 Perfect score: Title:

279 1 CTNCGTKTTPLWRRNPQGQP......LHGVVRPLSLKTDVIKKRQR 50 Sednence:

283224 segs, 96134422 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched:

Scoring table:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database

PIR\_73:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

|   | Description           | റ      | ator | цe   | probable nitrogen |      |      |      |        | transcription fact | Q,     |        |        |        | 44     | ₩.     |        |        |        | ion    | GATA-type transcri | ei     | BmGATA beta isofor | S      |        |      |      | 44     | Ęa   | transcription fact |
|---|-----------------------|--------|------|------|-------------------|------|------|------|--------|--------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------------------|--------|--------------------|--------|--------|------|------|--------|------|--------------------|
|   | ΙD                    | S51493 | 798  | 34   | 133               | 5623 | n    | 53   | I51420 | 151419             | JC6170 | S40382 | S22781 | I50702 | 150703 | A49645 | 150701 | A48099 | 157561 | 161183 | T38291             | S27473 | S53812             | S53811 | S04655 | ~    | 29   | A32993 | 488  | S48756             |
|   | DB                    | 1      |      |      |                   |      | 7    |      |        |                    |        |        |        |        |        |        |        |        |        |        |                    |        |                    |        |        |      |      |        |      |                    |
|   | Query<br>Match Length | 862    | 876  | 1036 | 855               | 510  | 730  | 551  | 388    | 390                | 532    | 779    | 269    | 391    | 387    | 391    | 380    | 439    | 439    | 440    | 564                | 950    | 241                | 327    | 413    | 509  | 564  | 304    | 413  | 413                |
| æ | Query<br>Match        |        |      | 91.8 |                   |      | 72.4 | 72.0 |        |                    |        |        |        |        | 67.4   | 67.4   | 66.3   | 66.3   | 66.3   | 66.3   | 62.9               | 65.6   | 63.8               | 63.8   | 63.8   | 63.8 | 63.8 | 63.4   | 63.4 | 63.4               |
|   | Score                 | 258    | 258  | 256  | 242               | 238  | 202  | 201  | 195    | 195                | 194    | 191    | 190    | 189    | 188    | 188    | 185    | 185    | 185    | 185    | 184                | 183    | 178                | 178    | 178    | 178  | 178  | 177    | 177  | 177                |
|   | Result<br>No.         |        | 7    | ٣    | 4                 | S    | 9    | 7    | 83     | 6                  | 10     | 11     | 12     | 13     | 14     | 15     | 16     | 17     | 18     | 19     | 20                 | 21     | 22                 | 23     | 24     | 25   | 26   | 27     | 28   |                    |

| n fact             | prote              | n fact        | n fact        | n fact        | n fact        |               |               |               |               |               |               |               | n fact        | n fact        | prote        |            | ,                  |
|--------------------|--------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------------|------------|--------------------|
| transcription fact | hypothetical prote | transcription | hypothetical |            |                    |
| A41267             | T26296             | A41602        | A57601        | C41602        | B48099        | B41602        | A36389        | 41782         | A40815        | D41602        | B39794        | A39794        | B36389        | A56953        | T19677       | ALIGNMENTS |                    |
| 1                  | 7<br>T             | 2<br>A        |               | -             |               |               | 2             |               |               |               | 2<br>E        | 1             | 2<br>E        | 2             | 2            |            |                    |
| 416                | 454                | 359           | 486           | 452           | 118           | 364           | 466           | 474           | 480           | 435           | 443           | 444           | 444           | 433           | 613          |            |                    |
| 63.4               | 63.4               | 62.7          | 62.7          | 62.4          | 62.0          | 62.0          | 62.0          | 62.0          | 62.0          | 61.6          | 61.6          | 61.6          | 61.6          | 55.4          | 55.4         |            |                    |
| 177                | 177                | 175           | 175           | 174           | 173           | 173           | 173           | 173           | 173           | 172           | 172           | 172           | 172           | 154.5         | 154.5        |            |                    |
| 30                 | 31                 | 32            | 33            | 34            | 35            | 36            | 37            | 38            | 39            | 40            | 41            | 42            | 43            | 44            | 45           |            | RESULT 1<br>S51493 |

```
Solidy:
Solidy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Introns: 123/2
S.Superfamily: nitrogen regulatory protein nit-2; GATA-type zinc finger homology
C;Keywords: zinc finger
F;659-712/Pomain: GATA-type zinc finger homology <G2F>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  662 CINCFIQITPLMRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 862;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 CINCGIKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.5%; Score 258; DB 2; I
illarity 92.0%; Pred. No. 6.7e-24;
Conservative 2; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Cross-references: EMBL:U02612
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 46; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-862 <HAA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
```

c. Species: Emericella nidulans C. Species: Emericella nidulans, Aspergillus nidulans C. Species: Emericella nidulans, Aspergillus nidulans C. Species: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 16-Jun-2000 C. Accession: A57988; Sloodly: S70168 ; S2683 R. Kiudla, B.; Caddick, M.X.; Langdon, T.; Martinez-Rossi, N.M.; Bennett, C.F.; Sibley, EMBO J. 9, 1355-1364, 1990 A.Title: The regulatory gene area mediating nitrogen metabolite repression in Aspergi A.Accession: A57988 A.Accession: A57988 A.Accession: A57988 A.Kolecule type: DNA A.Rolecule type: DNA A.Rolecule type: DNA

A;Cross-references: EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g1154625 A;Note: this sequence represents reinterpretation to include two exons A;Accession: S10017 A; Molecule type: DNA A; Residues: 158-876 <KU2> A;Cross-references: EMBL:X52491 A;Note: this sequence represents the authors' original translation

```
A;Reference number: Z21970
A;Accession: T41336
A;Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best_Local Similarity
Matches 41; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mar
Local Sim.
42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    φ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
S50543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
R;Langdon, T.; Sheerins, A.; Ravagnani, A.; Gielkens, M.; Caddick, M.X.; Arst Jr., H.N. Mol. Microbiol. 17, 877-888, 1995
A;Title: Mutational analysis reveals dispensability of the N-terminal region of the Aspe A; Reference number: S70167; MUID:96123430; PMID:8596437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; Species: Neurospora crassa
C;Date: 13-Jul-1990 #sequence_revision 26-Jul-1996 #text_change 16-Feb-1997
C;Accession: A34755
R;Fu, Y.H.; Marzluf, G.A.
Mol. Cell. Biol. 10, 1056-1065, 1990
A;Title: nit-2, the major nitrogen regulatory gene of Neurospora crassa, encodes a prote
A;Reference number: A34755; MUID:90158568; PMID:2137552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 probable nitrogen regulatory transcription factor gafl - fission yeast (Schizosaccharomy C;Species: Schizosaccharomyces pombe C;Date: 3C0-Oct-2000 #sequence_travision 08-Dec-2000 #text_change 08-Dec-2000 C;Accession: T41336; T41208; T42297 R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G. submitted to the EMBL Data Library, January 1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Superfamily: nitrogen regulatory protein nit-2; GATA-type zinc finger homology C;Reywords: DNA bluding; transcription regulation; zinc finger F;670-723/Domain: GATA-type zinc finger homology <GZF> F;673-697/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Superfamily: nitrogen regulatory protein nit-2; GATA-type zinc finger homology C;Keywords: DNA binding; transcription regulation; zinc finger F;740-793/Domain: GATA-type zinc finger homology <GZF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                     A;Molecule type: DNA
A;Rosidues: 1-876 <CAD>
A;Cross-references: EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g1154625
C;Genetics:
A;Gene: areA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 256; DB 1; Length 1036;
Pred. No. 1.4e-23;
1; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 258; DB 1; Length 876;
Pred. No. 6.8e-24;
2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         743 CINCFTQTTPLMRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                        A;Status: nucleic acid sequence not shown
A;Notecule type: DNA
A;Residues: 1-791, '4, '793, 'T', 795, 'SPGTNS', 802-876 <LAN>
A;Cross-references: EMBL:X52491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Description: mediates nitrogen metabolite repression
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nitrogen regulatory protein nit-2 - Neurospora crassa
                                                                                                                                                                                                                       R;Caddick, M.X. submitted to the EMBL Data Library, October 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;740-793/Domain: GATA-type zinc iluger F;743-767/Region: zinc finger GATA motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.5%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      91.8%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Cross-references: GB:M33956
                                                                                                                                                                                                                                                                       A; Reference number: S72883
A; Accession: S72883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: DNA; mRNA
A; Residues: 1-1036 <FUY>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
These 46; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Introns: 209/2; 335/3
                                                                                                   A; Accession: S70168
                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Introns: 147/2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C: Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
A; Status: Translated from GB/EMBL/DDBJ
A; Residues: 1-170 <LYN
A; Residues: 1-170 <LYN
A; Residues: 1-170 <LYN
A; Residues: 1-170 <LYN
A; Cross-references: EMBL:AL035076; NID:g4107277; PIDN:CAA22647.1; PID:g4107278; GSPDB
A; Cross-references: Extrain 972h-; cosmid c417
B; Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
A; Reseger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
A; Recession: T41208
A; Rocession: T41208
A; Rocession: T41208
A; Residues: 129-815 <AEE>
A; Cross-references: EMBL:AL049521; NID:g6318249; PIDN:CAB40003.1; PID:g4539591; GSPDB
A; Resedues: 129-815 <AEE>
A; Cross-references: EMBL:AL049521; NID:g6318249; PIDN:CAB40003.1; PID:g4539591; GSPDB
A; Reference number: 222401; MUID:98382525; PMID:9714831
A; Reference number: 222401; MUID:98382525; PMID:9714831
A; Reference number: 222401; MUID:98382525; PMID:9714831
A; Residues: Dreliminary; translated from GB/EMBL/DDBJ
A; Residues: 566-855 <AICCOSS-references: EMBL:L31601; NID:g710606; PIDN:AAC35593.1; PID:g710607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       probable membrane protein YFL021w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
C;Accession: S5623
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas submitted to the EMBL. Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccaromyces A;Reference number: S56186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Accession: S56233
A;Molecule type: DNA
A;Residues: 1-510 <MUR>
A;Cross-references: EMBL:D50617; NID:9836685; PID:d1009858; PID:9836733; MIPS:YFL021w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A Map position: 3R
C;Superfamily: GATA-type zinc finger homology
C;Keywords: transcription factor; transcription regulation; zinc finger
F;632-685/Domain: GATA-type zinc finger homology <GZF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            86.7%; Score 242; DB 2; Length 855; 84.0%; Pred. No. 6.3e-22; Live 5; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CINCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 CINCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A,Map position: 6L
C,Superfamily: GATA-type zinc finger homology
C,Keywords: transmembrane protein; zinc finger
E;14-30/Domain: transmembrane #status predicted <TMI>F;307-360/Domain: GATA-type zinc finger homology <G2F>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85.3%; Score 238; DB 2;
82.0%; Pred. No. 1.2e-21;
ilve 5; Mismatches 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Gene: SPDB:SPCC417.01c; SPDB:SPCC1902.01; gaf1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Gene: SGD:GAT1
A;Cross-references: SGD:S0001873; MIPS:YFL021w
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ilarity 84.0%;
Conservative
```

Mon Jan

```
GenCore version 5.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 19:42:57; Search time 48 Seconds (without alignments)
214.632 Million cell updates/sec
Title: 281
Sequence: 1 CINCFTQTTPLWRRNPDGQP......LHGVVRPLSLKTDVIKKRNR 50
Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580
Minimum DB seq length: 2000000000
```

Database : SPTREMBL\_21:\*
1: sp\_archea:\*
2: sp\_bacteria:\*
3: sp\_fungi:\*
4: sp\_human:\*
5: sp\_invertebrate:\*
6: sp\_nammal:\*
7: sp\_mhc:\*
8: sp\_crganelle:\*
9: sp\_phage:\*
10: sp\_ptin:\*
11: sp\_rodent:\*
11: sp\_rodent:\*
13: sp\_virus:\*
14: sp\_unclassified:\*
15: sp\_virus:\*
16: sp\_bacteriap:\*
17: sp\_archeap:\*

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

|           |   |        | Description     | Q9c414 aspergillus | Q9y7e8 aspergillus | Q9hew7 cladosporiu | 060043 metarhizium | Q9p952 schizophyll | P87016 zygosacchar | O59842 pericillium | Q9hev2 emericella | Q9hev3 emericella | Q9hev4 emericella | Q9w6u0 brachydanio | 077157 strongyloce | Q95vy5 anopheles q | Q8t3j0 drosophila | Q9nj18 ceratitis c | Q8wqq4 calliphora |
|-----------|---|--------|-----------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-------------------|-------------------|-------------------|--------------------|--------------------|--------------------|-------------------|--------------------|-------------------|
| SUMMARIES | • |        | ΠD              | Q9C414             | Q9Y7E8             | Q9HEW7             | 060043             | Q9P952             | P87016             | 059842             | Q9HEV2            | Q9HEV3            | Q9HEV4            | 09W6U              | 077157             | Q95VY5             | Q813J0            | Q9NJ18             | Q8WQQ4            |
|           |   |        | DB              |                    | ო                  | m                  | ٣                  | 3                  | ٣                  | e                  | e                 | ო                 | n                 | 13                 | S                  | Ŋ                  | S                 | S                  | 5                 |
|           |   |        | Match Length DB | 853                | 998                | 918                | 944                | 1075               | 327                | 298                | 312               | 320               | 436               | 383                | 119                | 537                | 486               | 324                | 511               |
|           | æ | Query  | Match           | 98.6               | 98.6               | 98.6               | 98.6               | 87.9               | 85.4               | 72.6               | 70.5              | 70.5              | 70.5              | 69.4               | 69.0               | 0.69               | 68.7              | 9.79               | 9.79              |
|           |   |        | Score           | 277                | 277                | 277                | 277                | 247                | 240                | 204                | 198               | 198               | 198               | 195                | 194                | 194                | 193               | 190                | 190               |
|           |   | Result | Q               | П                  | 7                  | æ                  | 4                  | 2                  | 9                  | 7                  | 80                | 6                 | 10                | 11                 | 12                 | 13                 | 14                | 15                 | 16                |

ö

Gaps

ö

QY Db

Query Match 98.6%; Score 277; DB 3; Length 853; Best Local Similarity 98.0%; Pred. No. 8.6e-30; Matches 49; Conservative 1; Mismatches 0; Indels

| 094884 drosophila 09vf01 drosophila 090410 brachydanio 09ptj2 brachydanio 09ptj2 brachydanio 09ptj1 brachydanio 09ptj1 brachydanio 09ptj1 brachydanio 077156 strongyloce 016365 home sapien 08w187 mus musculu 09w52 mus sp. ret 096w52 mus sp. ret 096w59 mus musculu 094w54 rattus norv 090ty1 brachydanio 090ty1 brachydania 090tf4 aedes aegyp 07426 neurosophila 090tf4 aedes aegyp 07426 neurosopnia 080tm6 encephalito 090tg1 drosophila | MENTS      | 853 AA.                            | Created)<br>Last sequence update)<br>Last annotation update) | AREA.<br>Aspergillus oryzae.<br>Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;<br>Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.<br>NCBI_TaxID=5062; | tool to overproduce a large number of   | EMBL/GenBank/DDBJ databases.<br>  | Interpro; IPR000609; Znf_GATA. Pfam, PF00320; GATA; 1. PRINTS; PR00619; GATAZNFINGER. SMART; SM00401; ZnF_GATA; 1. PROSITE; PS00344; GATA_ZN_FINGER_1; 1. PROSITE; PS50114; GATA_ZN_FINGER_2; 1. SEQUENCE 853 AA; 91497 MW; A5EC8ICE256A6B15 CRC64; |
|---|------------|------------------------------------|--|--|---|---|---|
| 094884<br>0994601<br>090410<br>099712<br>099713<br>099713<br>0907186<br>016365<br>016365<br>016365<br>0907187<br>0907187<br>0907187<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879<br>090879   | ALIGNMENTS | PRT;                               | Created)<br>Last sequ<br>Last anno                           | ; Pezizor<br>mitospor:   | a tool to   | EMBL/Gen]<br>   | A.<br>R.<br>NGER_1;<br>NGER_2;  |
| 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1   |            |                                    | 17,<br>17,<br>19,  | cota<br>ae;  | a   | the<br>.1;  | GAT<br>INGE<br>; 1.<br>N_FI<br>N_FI   |
| 0044880444114444414460466464646646466464  |            | PRELIMINARY;                       | (TrEMBLrel.<br>(TrEMBLrel.                                   | ASCOMY<br>COMACE   | A gene  | 0) to<br>K08066   | 9; Znf<br>A; 1.<br>ATAZNF<br>F_GATA<br>GATA_Z<br>GATA_Z<br>GATA_Z   |
| 297966831111187777778111158899966   |            | ELIM                               | remb<br>remb<br>remb   | yzae<br>gi;<br>icho<br>2;  | N.A.<br>P.;   | -200<br>; AA  | 444;<br>044;<br>144;<br>AA;<br>AA;  |
| 00000000000000000000000000000000000000  |            |                                    | 111  | Fun<br>Fun<br>Tr<br>=506   | ROM<br>;<br>oek<br>yzae   | (NOV<br>0305  | 1370;<br>320;<br>00061<br>0401<br>55003<br>853  |
| 11111111111111111111111111111111111111  |            | J 1<br>.4<br>Q9C414<br>Q9C414;     | 01-JUN-2001<br>01-JUN-2001<br>01-DEC-2001<br>AreA.           | AREA. Aspergillu Eukaryota; Eurotiales NCBI_TaxID  | Lin<br>SEQUENCE FROM N.A<br>STRAIN=TK3;<br>Van den Broek P.;<br>"The A. oryzae ar | proteases.";<br>Submitted (NOV-2000) to the EN<br>EMBL; AF320305; AAK08066.1; | terpro;<br>terpro;<br>am; PF00<br>INTS; PR<br>ART; SM0<br>OSITE; P<br>QUENCE  |
| 11112222222222222222222222222222222222  |            | RESULT<br>Q9C414<br>ID Q9<br>AC Q9 | A 0010   | ARE<br>ASP<br>Euk<br>Euk<br>NCB  | SEQU<br>STR   | Pr<br>Su<br>E   | SPRATE  |
|   |            | RES<br>Q9C<br>ID                   | DT<br>DE   | 000 000 000 000 000 000 000 000 000 00   | RP<br>RA<br>RA  | RT  | SORBRERS  |
|   |            |                                    |  |  |   |   |   |

```
PRT;
                                                                                                                                                                                                                                                                                                                MEDLINE=99070047; PubMed=9852945;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00320; GATA; 1.
PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZNF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Transcription factor ScGATA-6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98.6%;
98.0%;
                                                                                                                                                                                                      Nitrogen response regulator.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                49; Conservative
                           Conservative
                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
              Best Local Similarity
                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=5334;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yamagishi K.;
                                                                                                                                                                                                                                                                                                       STRAIN-ME1;
                           49;
   Query Match
                                                                                                                                          060043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9P952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09P952
                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                 RESULT 4
                                                                                                                           060043
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9P952
                                                     ŏ
                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .
0
                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN=ATCC 56775;
MEDLINE=20225467; PubMed=10760588;
Chang P.-K., Yu J., Bhatnagar D., Cleveland T.E.;
"Characterization of Aspergillus parasiticus major nitrogen regulatory
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
Perez-Garcia A., Snoeijers S.S., Joosten M.H.A.J., Goosen T.,
De Wit P.J.G.M.;
"Expression of the avirulence gene Avr9 of the fungal tomato pathogen
Cladosporium fulvum is regulated by the global nitrogen response
factor NRFI.";
                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                         Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBI_TaxID=5067;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-WAR-2001 (TrEMBLrel. 21, Last annotation update)
01-JUN -2002 (TrEMBLrel. 21, Last annotation update)
Nitrogen response factor NRF1.
Cladosporium fulvum (Fulvia fulva).
Eukaryota; Fungi; Ascomycota; Pezizomycotina;
Dothideomycetes et Chaetochyriomycetes incertae sedis;
Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
                                                                                                                                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 866;
                                                                                                                                                                                                                                                                                                                                                                                                                                        1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                  DISCRIPTO, FLITZ), FLITZ), FLITZ), FLITZ), FLITZ), FLITZ), FLITZ, FLOODSO, GATA, 1.

PRINTS; PROOG19; GATAZNITINGER.
SMART; SMO0401; ZAF_GATA; 1.

PROSITE; PS50144; GATA_ZN FINGER_1; 1.

PROSITE; PS50114; GATA_ZN FINGER_2; 1.

SEQUENCE 866 AA; 92877 MW; 16FFF4D56E823988 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00320; GATA; 1.

PRINTS; PR00619; GATAZNFINGER.

PRINTS; PR01217; PRICHEXTENSN.

SMART; SM004011; ZnE_GATA; 1.

PROSITE; PS00344; GATA_ZN_FINGER_1; 1.

PROSITE; PS50114; GATA_ZN_FINGER_2; 1.

SEQUENCE 918 AA; 99077 MW; E98198D99BCA899 CRC64;
                                              01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                        98.6%; Score 277; DB 3;
98.0%; Pred. No. 8.7e-30;
iive 1; Mismatches 0
                        866 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                918 AA
                                                                                                                                                                                                                                       gene, area.";
Blochim. Blophys. Acta 1491:263-266(2000)
EMBL; AF148539; AAD37409.1;
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mol. Plant Microbe Interact. 0:0-0(2001).
EMBL; AF312694; AAG48616.1; -.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002965; P_rich_extensn.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                    Major nitrogen regulatory protein.
                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 98.03
Matches 49; Conservative
                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                             Aspergillus parasiticus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=5499;
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9HEW7;
                        Q9Y7E8
                                     Q9Y7E8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9HEW7
RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09HEW7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
```

```
ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Metarhizium anisopliae.
Eukaryota; Fungl; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae; Metarhizium.
NCBI_TaxID=5530;
                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Schizophyllaceae; Schizophyllum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Screen S.E., Bailey A.M., Charnley K., Cooper R., Clarkson J.M.;
"Isolation of a nitrogen response regulator gene (nrrl) from Metarhizium anisopliae.";
Metarhizium anisopliae.";
EMBL: 411999.
EMBL: AJ006468; CAA07052.1;
HSSP: P17429; 4GAT.
                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 277; DB 3; Length 944;
Pred. No. 9.6e-30;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             663 CINCFIQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 712
                                                                                                                                                     673 CTNCFTQTTPLMRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 722
Length 918;
                                                                                                                     1 CINCFIOTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "putative transcription factor ScGATA-6.";
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AB043467; BAA96108.1; -.
HSSP; P17429; 4GAT.
InterPro; IPR000679; 2nf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HSSE, TITALS, "ALGATA.
INTERPROPOSED 201_GATA.
Pfam; PF00320; GATA; 1.
PRINTS; PR00619; GATAZNITNGER.
SMART; SM00401; ZDF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
SEQUENCE 944 AA; 99687 MW; A1A723E658C23EA0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                            01-AUG-1998 (TrEMBLrel. 07, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                            ö
98.6%; Score 277; DB 3; 98.0%; Pred. No. 9.3e-30;
                                                            1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Schizophyllum commune (Bracket fungus).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-077-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequ
01-JUN-2002 (TrEMBLrel. 21, Last anno
```

Mon Jan

1

```
PRINTS; PR00619; GATAZNFINGER.
                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9HEV3;
01-MAR-2001 (
01-MAR-2001 (
                                                                  Query Match
                                                                                                                                                                                         Q9HEV2
                                                                                                                                                                                09HEV2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9HEV3
                                                                                      Matches
                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                                           RESULT
                                                                                                                                                                      09HEV2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9HEV3
       SOR
                                                                                                                                                                                          δλ
                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ID
DT
DT
DE
OC
                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                 Gaps
                                                                                                                                                                                                                                                                                                              "Cloning of the SAT1 gene concerned with salt tolerance of the yeast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
                                                                                                                                                                                                                                         Eukaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes, Saccharomycetales, Saccharomycetaceae, Zygosaccharomyces.
                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                            Ushio K., Otsuka H., Yoshikawa S., Taguchi G., Shimosaka M.
Mitsui N., Okazaki M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 327;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
PS00344; GATA_ZN_FINGER_1; 1.
PS50114; GATA_ZN_FINGER_2; 1.
1075 AA; 113322 MW; D29343211C02FFEA CRC64;
                                                               4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Haas H., Angermayr K., Zadra I.;
Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
EMBL: U96385; AaC09045.1; -.
HSSP; P17429; 4GAT.
InterPro; IRR006679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                DNA-binding.
SEOUENCE 327 AA; 37304 MW; 9D7AF1002807283F CRC64;
                                                                                                                                                                           (TrEMBLrel. 04, Created)
(TrEMBLrel. 04, Last sequence update)
(TrEMBLrel. 17, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TrEMBLrel. 07, Last sequence update) (TrEMBLrel. 17, Last annotation update)
                                          Score 247; DB 3;
Pred. No. 1.6e-25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 85.4%; Score 240; DB 3;
86.0%; Pred. No. 4.3e-25;
ilve 3; Mismatches 4;
                                                                                                                                                         327 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 298 AA.
                                                             2; Mismatches
                                                                                                                                                                                                                              Zygosaccharomyces rouxii (Candida mogii).
                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1998 (TrEMBLrel. 07, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                        Zygosaccharomyces rouxii.";
2. Ferment. Bloeng. 82:16-21(1996).
EMBL; D93211; BAA18948.1;
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000679; Znf_GATA, Pfam; PF00320; GATA; 1. PRINTS; PR00619; GATAZNPINGER. SMART; SM00401; ZnF_GATA; 1.
                                                   88.08;
                                           84.78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GATA transcription factor.
                                                              44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Penicillium chrysogenum
                                                                                                                                                                                                           protein.
                                                    Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                             NCBI_TaxID=4956;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=5076;
                                                                                                                                                                            01-JUL-1997
                                                                                                                                                                                                          DNA-binding
                                                                                                                                                                                       01-JUL-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1998
01-JUN-2001
                                                                                                                                                                                                01-JUN-2001
                      SEQUENCE
  PROSITE;
              PROSITE;
                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                      P87016
P87016;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          059842
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               059842
                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                  RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
                                                      Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE=89218944; PubMed=2651886;
Arst H.N. Jr., Tollervey D., Caddick M.X.;
"A translocation associated, loss-of-function mutation in the nitrogen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      metabolite repression regulatory gene of Aspergillus nidulans can revert intracistronically.";
                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Emericella nidulans (Aspergillus nidulans).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; Emericella.
NCBI_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Emericella nidulans (Aspergillus nidulans).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
                                                                                                      Length 298;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 312;
                                                                                                                                                                                  1 CINCFIQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                         99
                                                                                                                                              10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
SMART; SM00401; ZNF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
SEQUENCE PS9114; A374, 32670 MW; C220D36F4993C590 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIAM; Frussu; Gararinger.
PRINTS: PRO0619; GATARINGER.
SMART; SMO401; ZnE_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
SEOUENCE 312 AA; 34279 MW; 35A73B7526FE5F88 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                    72.6%; Score 204; DB 3; 76.0%; Pred. No. 3.9e-20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 198; DB 3;
Pred. No. 2.8e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         70.5%; Scor.
74.0%; Pred. No. 2...
2; Mismatches
                                                                                                                                                                                                                                                                                                                             312 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A
                                                                                                                                          2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mol. Gen. Genet. 215:364-367(1989),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Interpro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                              01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-JUN-2002 (TrEMBLrel. 21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2001 (TrEMBLrel. 16,
01-MAR-2001 (TrEMBLrel. 16,
01-UTM-2002 (TrEMBLrel. 21,
GATA factor AREB beta.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 74.0°
Matches 37; Conservative
                                                                                                                                            38; Conservative
                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                           GATA factor AREB alpha.
```

SEOUENCE

```
436 AA; 46251 MW; F91457AE2F8F7BB6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=99328904; PubMed=10398804;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 64.0
Matches 32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Strongylocentrotus.
                                  Best_Local Similarity
Matches 37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=7668;
                                                                                                                                                                                                                                             GATAS OR GTAS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            077157;
                                                                                                                                                                   009W60
                                                                                                                                                                                009W60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              077157
                                                                                                                                                                                                                                  Gata5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                         RESULT 11
                                                                                                                                                       009W60
                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AC OC OC OC OX NX RP RP RX
                                                                              δλ
                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-89218944; PubMed-2651886;
Arst H.N. Jr., Tollervey D., Caddick M.X.;
Arst Holorion associated, loss-of-function mutation in the nitrogen metabolite repression required procession required function associated.

Mol. Gen. Genet. 215:364-367(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE-21206224; PubMed-11309119;
MEDLINE-21206224; PubMed-11309119;
Conlon H., Zadra I., Haas H., Arst H.N. Jr., Jones M.G., Caddick M.X.;
The Aspergillus nidulans GATA transcription factor gene areB encodes at least three proteins and features three classes of mutation.";
Mol. Microbiol. 40:361-375(2001).
HSSP; P17429; 4GAT.
                                      SEQUENCE FROM N.A.
MEDLINE-89218944; Pubmed-2651886;
Arst H.N. Jr., Tollervey D., Caddick M.X.;
"A translocation associated, loss-of-function mutation in the nitrogen metabolite repression regulatory gene of Aspergillus nidulans can
                                                                                                                                         SEQUENCE FROM N.A. MEDLINE-21206224; PubMed=11309119; MEDLINE-21206224; PubMed=11309119; Conlon H., Zadra I., Haas H., Arst H.N. Jr., Jones M.G., Caddick M.X.; "The Aspergillus nidulans GATA transcription factor gene areB encodes at least three proteins and features three classes of mutation."; Mol. Microbiol. 40:361-375(2001). EMBL: AF320976; AAG49352.1; -- HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Emericella nidulans (Aspergillus nidulans).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
                                                                                                                                                                                                                                                                                                                                              70.5%; Score 198; DB 3; Length 320; 74.0%; Pred. No. 2.9e-19; tive 2; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                   1 CINCFIQTIPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                               25 CONCGTSKTPLWRRDELGSVLCNACGLFLKLHGRPRPISLKTDVIKSRNR 74
                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRINTS; PRO0619; GATAZNFINGER.
SMART: SMO0401; ZNE_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
NON_TER
  Eurotiales; Trichocomaceae; Emericella.
NCBI_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eurotiales; Trichocomaceae; Emericella
                                                                                                      revert intracistronically.";
Mol. Gen. Genet. 215:364-367(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GATA factor AREB gamma (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR000679; Znf GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 74.03
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9HEV4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q9HEV4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AREB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09HEV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
ö
                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Rodaway A., Takeda H., Koshida S., Broadbent J., Price B.M.J., Smith J.C., Patient R., Holder N.; Smith J.C., Patient R., Holder N.; Induction of the mesendoderm in the zebrafish germ ring by yolk cell-derived TGF-(beta) family signals and discrimination of mesoderm and endoderm by FGF."; Development 126:1067-3077(1999).

EMBL: AJ242515; CAB43400.1; -- HSSP; P17679; IGNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Strongylocentrotus purpuratus (Purple sea urchin).
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Brachydanio rerio (Zebrafish) (Zebra danio).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
0
                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 383;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241 CINCHISTITHERNAEGEPVCNACGLYMKLHGVPRPLAMKKESIQTRKR 290
                                                                                                                                                 DB 3; Length 436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                     20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             69.4%; Score 195; DB 13; Length 36 64.0%; Pred. No. 9.1e-19;
                                                          11; Indels
                                                                                                                        1 CINCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZFIN ZDB-GRNE-980256-340; gata5.
ZFIN ZDB-GRNE-980256-340; gata5.
InterPro; IPR001164; hRIP_like.
InterPro; IPR001679; Znf_GATA.
Pfam; PF00320; GATAZNFINGER.
SMART; SM00401; ZnF_GATA; 2.
PROSITE; PS00144; GATA_ZN_FINGER_1; 2.
PROSITE; PS00114; GATA_ZN_FINGER_1; 2.
SROSITE; PS0114; GATA_ZN_FINGER_2; 1.
SEQUENCE 383 AA; 41479 MW; 7F752664185588AF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-NAR-2002 (TrEMBLrel. 20, Last annotation update)
GATA transcription factor (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
   Score 198; DB 3;
Pred. No. 4e-19;
70.5%; Score 198; DB
nilarity 74.0%; Pred. No. 4e-1
Conservative 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TISSUE=EMBRYO;
MEDLINE=99307072; PubMed=10375499;
```

```
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                            SEQUENCE FROM N.A. Stapleton P., Hong L., Agbayani A., Carlson J., Stapleton M., Brokstein P., Hong L., Dresnek D., Farfan D., Frise E., Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G., Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S., Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
                                                                                                                                                                                                                                                                         Celliker S.;
Celliker S.;
Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
EMBL: AY095095; AAM11423.1; -.
SEQUENCE 486 AA; 51606 MW; BCOE10E9BBD40A0B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wuelbeck C., Simpson P.;
Wuelbeck C., Simpson P.;
The expression of pannier and achaete-scute homologs in a mosquito suggests an ancient role of pannier as a selector in the regulation of the dorsal body pattern.";
Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF395080; AAK97462.1;
InterPro; IPR001164; hRIP_like.
InterPro; IPR001164; pri_GATA.
                                        transcription factors and homologues of effector
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zinc finger transcription factor pannier.
Anopheles gambiae (African malaria mosquito).
Bukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                          genes of the vertebrate immune system expressed in sea urchin coelomocytes.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         69.0%; Score 194; DB 5; Length 119; 66.0%; Pred. No. 3.7e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 537;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9; Indels
                                                                                                                                                                                                                                                                                                                                        119 119
119 AA; 13455 MW; D924DA417C2B42EC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00320; GATA; 2.
PROSITE; PS00344; GATA_ZN_FINGER_1; UNKNOWN_2.
PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
SEQUENCE 537 AA; 55906 MW; A25BCOCF53EF17B0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (TrEMBLrel. 19, Created)
(TrEMBLrel. 19, Last sequence update)
(TrEMBLrel. 20, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         69.0%; Score 194; DB 5;
66.0%; Pred. No. 1.8e-18;
1ve 8; Mismatches 9
                                                                                                                                                                                                                                                                                                                                                                                                                        66.0%; Pred. ....
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           537 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         486 AA.
                                                                                                                                                                                                                            PRINTS, PROOGIS; GATAZNITNGER.
SMART; SMO0401; ZnF_GATA; 2.
PROSITE; PSO0344; GATA_ZN_FINGER_1; 2.
PROSITE; PSS0114; GATA_ZN_FINGER_2; 2.
             Pancer Z., Rast J.P., Davidson E.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                               Immunogenetics 49:773-786(1999).
EMBL; AF077675; AAC62961.1; -.
HSSP; P17679; IGNF.
InterPro; IPR0001164; hRIP_like.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel. 21, C
(TrEMBLrel. 21, I
(TrEMBLrel. 21, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                     33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                    Origins of immunity:
                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=7165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-JUN-2002 (
01-JUN-2002 (
01-JUN-2002 (
SD02611P.
PNR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001
01-MAR-2002
                                                                                                                                                                                                                                                                                                                                             NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                         NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q8T3J0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        095VY5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q8T3J0
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               095VY5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QRI3JO
                  RAT RET RET DE R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
```

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wulbeck C., Simpson P.;
"Expression of achaete-scute homologues in discrete proneural clusters on the developing notum of the medfly Ceratitis capitata, suggests a common origin for the stereotyped bristle patterns of higher
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ceratitis capitata (Mediterranean fruit fly).
Eukaryota, Metazoa, Arthropoda, Tracheata; Hexapoda; Insecta;
Pherygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Tephritoidea; Tephritidae; Ceratitis.
                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
68.7%; Score 193; DB 5; Length 486; 66.0%; Pred. No. 2.2e-18; 1ive 8; Mismatches 9; Indels
                                                                                                                             Whelber C., Simpson P.;
Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AR184154; AAF61481.1; -.
HSSP; P17679; LORN.
InterPro; IPR001164; hRIP_like.
InterPro; IPR00679; Znf_GATA.
Pfam: PF001220; GATA.ZnFINGER.
SMART; SM00401; ZnF_GATA.
PROSITE; PS00134; GATA_ZN_FINGER_1; Z.
PROSITE; PS00134; GATA_ZN_FINGER_2; Z.
PROSITE; PS00134; GATA_ZN_FINGER_2; Z.
SEQUENCE 324 AA; 34785 MW; CFCE5F9D3CBD290B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          67.6%; Score 190; DB 5; Length 324; 64.0%; Pred. No. 3.8e-18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                 1 CTNCFTQTTPLWRRNPDGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR
                                                                                                                                                                                                                                                                                                                   01-077-2000 (TrEMBLrel. 15, Created)
01-077-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
2inc finger DNA binding protein pannier.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64.0%; Preu. ...
                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=BENAKEO;
MEDLINE=20171265; PubMed=10704387;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Development 127:1411-1420(2000).
                       Local Similarity 66.0 es 33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32; Conservative
                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=7213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-BENAKEO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                        Q9NJ18
Q9NJ18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Diptera.
                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                   Q9NJ18
```

Search completed: January 3, 2003, 19:50:35 Job time: 49 secs

```
GenCore version 5.1.3 Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

January 3, 2003, 18:50:12; Search time 14 Seconds (without alignments) 148.130 Million cell updates/sec Run on:

Title: Perfect score:

US-09-725-010-2 279 1 CTNCGTKTTPLWRRNPQGQP......LHGVVRPLSLKTDVIKKRQR 50 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

112892 segs, 41476328 residues Searched:

112892 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_40:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | Description    |            | 582        |            |            |            |            |            |            |            |            | P43574 saccharomyc | Q9hev5 neurospora | _          |            | drosophi  |            |            | Q92259 penicillium | _         |            | wns wnsc   | xenobns    |            | homo sa    | 3 gallus g | xenobns    | Q91677 xenopus lae | P46153 rattus norv | Q61169 mus musculu | 3 homo sa  | gal.       | mus mus    | 100 Cathery Cathery |
|-----------|----------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|--------------------|-------------------|------------|------------|-----------|------------|------------|--------------------|-----------|------------|------------|------------|------------|------------|------------|------------|--------------------|--------------------|--------------------|------------|------------|------------|---------------------|
|           | ID             | NUT1_MAGGR | AREA_PENCH | AREA_PENRO | NRFA_PENUR | AREA_ASPOR | AREA_EMENI | AREA_ASPNG | AREA_GIBFU | NIT2_NEUCR | GAF1_SCHPO | GAT1_YEAST         | ASD4_NEUCR        | GLN3_YEAST | GZF3_YEAST | PNR_DROME | GA5B_XENLA | GASA_XENLA | SREP_PENCH         | SRP_DROME | DA80_YEAST | GAT5_MOUSE | GA6A_XENLA | GAT5_CHICK | GAT5_HUMAN |            | GA6B_XENLA | GAT4_XENLA         | GAT6_RAT           | GAT6_MOUSE         | GAT6_HUMAN | GAT4_CHICK | GAT4_MOUSE | EAG VEAC            |
|           | DB             | П          |            |            |            |            |            |            |            |            |            |                    |                   |            |            |           |            |            |                    |           |            |            |            |            |            |            |            |                    |                    |                    |            | -          |            |                     |
|           | Length         | 926        | 725        | 860        | 865        | 998        | 876        | 882        | 971        | 1036       | 855        | 510                | 426               | 730        | 551        | 540       | 388        | 390        | 532                | 779       | 269        | 404        | 391        | 391        | 397        | 387        | 391        | 392                | 441                | 444                | 449        | 380        | 440        | 077                 |
| đ         | Query<br>Match | 93.5       | 92.2       | 92.2       | 92.5       | 92.5       | 92.5       | 92.5       | 92.5       | 91.8       | 86.7       | 85.3               | 75.3              | 72.4       | 72.0       | 71.0      | 6.69       | 6.69       | 69.5               | 68.5      | 68.1       | 68.1       | 67.7       | 67.7       | 67.7       | 67.4       | 67.4       | 67.4               | 67.4               | 67.4               |            | 66.3       |            |                     |
|           | Score          | 7          | 258        | 258        | 258        | 258        | 258        | 258        | 258        | 256        | 242        | 238                | 210               | 202        | 201        | 198       | 195        | 195        | 194                | 191       | 190        | 190        | 189        | 189        | 189        | 188        | 188        | 188                | 188                | 188                | 188        | 185        | 185        | 185                 |
|           | Result<br>No.  |            | 7          | 9          | 4          | S          | 9          | 7          | 80         | σ          | 10         | 11                 | 12                | 13         | 14         | 15        | 16         | 17         | 18                 | 19        | 20         | 21         | 22         | 23         | 24         | 25         | 26         | 27                 | 28                 | 29                 | 30         | 31         | 32         | 33                  |

| P43694 homo sapien | P40349 ustilago ma | P17679 mus musculu | P17678 gallus gall | P15976 homo sapien | P43429 rattus norv | P28515 caenorhabdi | P23767 xenopus lae | P91623 drosophila | P23770 xenopus lae |  |
|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-------------------|--------------------|--|
|                    |                    |                    |                    |                    |                    |                    |                    |                   |                    |  |
| GAT4_HUMAN         | URB1_USTMA         | GAT1_MOUSE         | GATI_CHICK         | GAT1_HUMAN         | GAT1_RAT           | ELT1_CAEEL         | GAIA_XENLA         | GATC_DROME        | GAT2_XENLA         |  |
|                    |                    | ٦.                 |                    | 7                  | Н                  | П                  | Н                  | Н                 | Н                  |  |
| 442                | 950                | 413                | 304                | 413                | 413                | 416                | 359                | 486               | 452                |  |
| 66.3               | 65.6               | 63.8               | 63.4               | 63.4               | 63.4               | 63.4               | 62.7               | 62.7              | 62.4               |  |
| 185                | 183                | 178                | 177                | 177                | 177                | 177                | 175                | 175               | 174                |  |
| 34                 | 36                 | 37                 | 20 ON<br>17 ON     | 40                 | 41                 | 42                 | 43                 | 44                | 45                 |  |

# ALIGNMENTS

```
FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIMIS; PROOF19; GATACAFINGER.
SMART; SMOO401; ZNF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nitrogen regulatory protein NRFA.
                                                                                                                                         EMBL; AJ001530; CAA04815.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                               InterPro; IPR000679; Znf_GATA. Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                       91615 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSFAC; T02828; -.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; U53137; AAB17740.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                          Best Local Similarity 92.0
Matches 46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF00320; GATA;
                                                                                                                                                                                                                                                                         860 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Penicillium urticae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. STRAIN-NRRL 2159A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=29844;
                                                                                                                                                                                                                                                                                                                                                                                                                               NRFA_PENUR
092269;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ellis C.M.;
                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                               Query Match.
                                                                                                                                                                                                                                                                                                                                                                                                                   NRFA_PENUR
ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                    Gaps
                                                                                                                          Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
NCBI_TaxID=5076;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
NCBI_TaxID=5082;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Transcription regulation, Activator, DNA-binding, Zinc-finger; Nuclear protein; Nitrate assimilation. 2N_FING 525 549 Gapa-mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.5%; Score 25%; DB 1; Length 725; 92.0%; Pred. No. 6e-25; Live 2; Mismatches 2; Indels
                                                    01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Nitrogen regulatory protein area (Nitrogen regulator nre).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKIDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-DEC-1998 (Rel. 37, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Nitrogen regulatory protein areA (Nitrogen regulator nmc).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gente S., Poussereau N., Fevre M.;
Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        45C24148F81F8D43 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                860 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00401; ZnF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
PRINTS; PR00619; GATAZNFINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Rel. 36, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      725 AA; 76848 MW;
                                                                                                                                                                                                                                                                                                                                                                                         EMBL; U02612; AAA83400.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      46; Conservative
                                                                                                               Penicillium chrysogenum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Penicillium roqueforti.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                      HSSP; P17429;
                                                                                                    AREA OR NRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-JUL-1998
                             AREA_PENCH
Q01582;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PENRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AREA OR NMC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-P2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                   AREA_PENCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AREA_PENRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AREA_
        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AC
DT
DT
DT
DT
DT
CO
OC
OC
OC
OC
NR
RR
RR
RR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota, Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes,
Eurotiales, Trichocomaceae, mitosporic Trichocomaceae, Penicillium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Transcription regulation; Activator; DNA-binding; Zinc-finger;
Nuclear protein; Nitrate assimilation
ZN FING 660 684 GATA-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92.5%; Score 258; DB 1; Length 860; 92.0%; Pred. No. 7e-25; ive 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     660 CINCFIQTIPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2B8BF0DDCDC1BA68 CRC64;
-!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Thesis (1996), University of Calgary, Canada.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRINTS: PRODE19; GATAZNFINGER.
SMART; SMO0401; ZNF_GATA; 1.
PROSITE; PSO0344; GATA_ZN_FINGER_1; 1.
PROSITE: PS50114; GATA_ZN_FINGER_2; 1.
```

```
Emericella nidulans (Aspergillus nidulans).
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
  16-ocT-2001 (Rel. 40, Last annotation update)
Nitrogen regulatory protein areA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dp
    DDD BRYNDD DRYNDD DRYND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL Outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ő
                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Christensen T., Hynes M.J., Davis M.A.;
submitted (NOV-1997) to the EMBL/Genbank/DDBJ databases.
-!- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN. POSITIVELY ACTING
REGULATORY GENE OF NITROGEN METABOLITE REPRESSION (BY SIMILARITY).
-!- SUBCELLUIAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRINTS; PRODGIG; GATAZNETNGER.
SMART; SMO0401; ZAF_GATA, 1.
PROSTET; PRO0344; GATA_ZN_FINGER_1; 1.
PROSTET; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
    DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  92.5%; Score 258; DB 1; Length 866; 92.0%; Pred. No. 7e-25; Live 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    665 CINCFIQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 714
                                                                                                           Length 865;
                                                                                                                                                                                             1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                   2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GATA-TYPE.
C09A03EB12E3FBB4 CRC64;
                                                                CC901BED187D8EEF CRC64;
                                                                                                           DB 1;
                                                                                                                                                                                                                                                                                                                                                                                  15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                               Pred. No. 7e-25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            876 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Rel. 15, Created)
(Rel. 33, Last sequence update)
                                                                                                                                                2; Mismatches
                                                                                                             Score 258;
                                            GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nuclear protein; Nitrate assimilation.
Transcription regulation; Activator; D
Nuclear protein; Nitrate assimilation.
ZN_FING 665 689 GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nitrogen regulatory protein areA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AJ002968; CAA05776.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; P17429; 4GAT.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                              865 AA; 92407 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         866 AA; 92882 MW;
                                                                                                        92.5%;
92.0%;
                                                                                                                                                     46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Aspergillus oryzae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-IFO 4177;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=5062;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-AUG-1990
01-FEB-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AREA_EMENI
P17429;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46;
                                                                                                                                                                                                                                                                                                                                              AREA_ASPOR
                                                                SEQUENCE
                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AREA_EMENI
                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                        AREA_ASPOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
                                                                                                                                                                                                                                                                                                                                              PTACE
    KW
KW
FT
                                                                                                                                                                                             δŽ
                                                                                                                                                                                                                                     q
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                        MEDLINE-90228331; PubMed-1970293, Kudla B., Caddick M.X., Langdon T., Martinez-Rossi N.M., Bennett C.F., Sibley S., Davies R.W., Arst H.N. Jr.; Tergeniatory gene areA mediating nitrogen metabolite repression in Aspergillus nidulans. Mutations affecting specificity of gene activation alter a loop residue of a putative zinc finger."; EMBO J. 9:1355-1364(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                           Langdon T., Secrins A., Ravagnani A., Caddick M.X., Arst H.N. Jr.; "Mutational analysis reveals dispensability of the N-terminal region of the Aspergillus transcription factor mediating nitrogen metabolite repression.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      factors.";
J. mol. mol. 277:605-620(1998).
-!- FUNCTION: TRANSCRIPTION ACTIVATOR THAT MEDIATES NITROGEN
METABOLITE REPRESSION IN A.NIDULANS. ACTIVATE THE TRANSCRIPTION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "The solution structure of a fungal AREA protein-DNA complex: an alternative binding mode for the basic carboxyl tail of GATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS00344; CATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   673 CINCFIQITPLWRRNPEGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRNR 722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 258; DB 1; Length 87
Pred. No. 7.1e-25;
2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE=98202574; PubMed=9533883;
Starich M.R., Wikstroem M., Arst H.N. Jr., Clore G.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      742 H-T-H MOTIF (PROBABLE).
94195 MW; 9ADC2273EE536F98 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nuclear protein; Nitrate assimilation; 3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
Eukaryota; Fungi; Ascomycota; Pezizomyc
Eurotiales; Trichocomaceae; Emericella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBCELLULAR LOCATION: Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mol. Microbiol. 17:877-888(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRINTS; PR00619; GATAZNFINGER. SMART; SM00401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000679; Znf_GATA. Pfam, PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.5%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; X52491; CAA36731.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [3]
STRUCTURE BY NMR OF 662-727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB; 4GAT; 28-JAN-98.
PDB; 5GAT; 28-JAN-98.
PDB; 6GAT; 28-JAN-98.
PDB; 7GAT; 28-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       876 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; S10017; S10017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSFAC; T02533;
                                                               NCBI_TaxID=5072;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gronenborn A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                    REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ZN_FING
DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
```

ô

```
PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIT2_NEUCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
ID DTT PD DTT PD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gibberella fujikuroi (Bakanae and foot rot disease fungus) (Fusarium
                                                                                                                                                                                                                                                       Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
NCBL_TaxID=5061;
                                                                                                                                                                                                                                                                                                                                                                                 Maccabe A.P., Vanhanen S.A.S., Sollewijn Gelpke M.,
van de Vondervoort P., Arst H.N., Visser J.;
Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN. POSITIVELY ACTING
REGULATORY GENE OF NITROGEN METABOLITE REPRESSION.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
Nuclear protein; Nitrate assimilation.
ZN_FING 676 700 GATA_TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota: Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreales; Nectriaceae; Gibberella.
NCBL_TaxID=5127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92.5%; Score 258; DB 1; Length 882; 92.0%; Pred. No. 7.1e-25; Indels iive 2; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97A502936B94E5AE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Rel. 36, Created)
(Rel. 36, Last sequence update)
(Rel. 41, Last annotation update)
                                                                                                                                            15-JUL-1998 (Rel. 36, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                 882 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GATA-TYPE
                                                                                                                                                                                     Nitrogen regulatory protein area.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nitrogen regulatory protein areA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FIAM; FF00320; GATA; 1.
PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZNF_GATA; 1.
                                                                                                                          15-JUL-1998 (Rel. 36, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               882 AA; 94518 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X81998; CAA57524.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                     Aspergillus niger.
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. STRAIN=m567;
                                                                                                                                                                                                                                                                                                                                                                    STRAIN-N402;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              moniliforme)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15-JUL-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-JUN-2002
                                                                               AREA_ASPNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AREA_GIBFU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                      013412;
                                                          AREA_ASPNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AREA_GIBFU
                                                                                                                                                                                                                AREA.
                                    RESULT 7
                                                                                                      ó
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                   Tudzynski B., Homann V., Feng B., Marzluf G.A.;
"Isolation, characterization and disruption of the area nitrogen
regulatory gene of Glabberella fujikurol.";
Mol. Genet. 261:106-114(1999).
-! FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN. POSITIVELY ACTING
REGULATORY GENE OF NITROGEN METABOLITE REPRESSION (BY SIMILARITY).
-! SUBCELLULAR LOCATION: Nuclear.
-! SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fu Y.-H., Marzluf G.A.; "Site-directed mutagenesis of the 'zinc finger' DNA-binding domain of the nitrogen regulatory protein NIT2 of Neurospora."; Mol. Microbiol. 4:1847-1852(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OF NITROGEN LIMITATION IT TURNS ON THE EXPRESSION OF GENES FOR ENZYMES WHICH ARE REQUIRED FOR THE USE OF A VARIETY OF SECONDARY NITROGEN SOURCES, INCLUDING NITRATES, PURINES, AMINO ACIDS, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-90158568; PubMed-2137552; Fu Y.-H., Marzluf G.A.; "nit-2, the major nitrogen regulatory gene of Neurospora crassa, encodes a protein with a putative zinc finger DNA-binding domain."; Mol. Cell. Biol. 10:1056-1065(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Interpro; IRR000679; Znf_GATA.
PFAM; PP00320; GATA; 1.
PRINTS; PR00619; GATAZNEINGER.
SMART; SM00401; ZnF_GATA_I.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS0114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
Nuclear protein; Nitrate assimilation.
ZN_FING

SOURTH STANDA SOURTH S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
NCBL_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 258; DB 1; LL. 2.
Pred. No. 7.7e-25; Lindels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  971 AA; 103580 MW; 887DD882141C7453 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1990 (Rel. 16, Created)
01-NOV-1990 (Rel. 16, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nitrogen catabolic enzyme regulatory protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 1036 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2; Mismatches
MEDLINE=99168774; PubMed=10071216;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-91186820; PubMed-2150539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; Y11006; CAA71897.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92.5%;
92.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4GAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN=74-OR23-1A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Neurospora crassa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P17429;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MUTAGENESIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NIT2_NEUCR
```

Mon Jan

```
GAT1_YEAST
P43574;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GAT1_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
        RRARE RRAPA PRAPA 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OC OS OC OC OC
                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-21848401; PubMed=11859360;
Mood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., Janes K., Jones L., Leather S., McDonald S., McLean J., Mooney P., Moule S., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                     PRINTS; PRUVULS, SMARN; SMARN; SMORUT; SMORUT; SMORUT; PROSITE; PS00044; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNa-binding; Zinc-finger; Nuclear protein; Nitrate assimilation; Repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GATA-TYPE. CTNC->STNG: ABOLISHES DNA-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 256; DB 1; Length 1036;
Pred. No. 1.5e-24;
1; Mismatches 3; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RR->GG: ABOLISHES DNA-BINDING.
NA->DV: ABOLISHES DNA-BINDING.
GL->DV: ABOLISHES DNA-BINDING.
KR->NS: ABOLISHES DNA-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5FE4992B1C223514 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                    INDUCTION: BY LACK OF A PRIMARY NITROGEN SOURCE.
                                        -!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GAF1 OR SPCC417.01C OR SPCC1902.01.
  -!- SUBCELLULAR LOCATION: Nuclear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      109310 MW;
                                                                                                                                                                                                                                                                                        TRANSFAC; T00627;
InterProc; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
PRINTS; PR00619; GATAZNFINGER.
                                                                                                                                                                                                                               EMBL; M33956; AAB03891.1; -.
PIR; A34755; A34755.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ch 91.8%;
1 Similarity 92.0%;
46; Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55
1110
767
746
756
769
790
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Schizosaccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN=972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MUTAGEN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
REPEAT
ZN_FING
MUTAGEN
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BWED cutstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
Weltjens I., Vanstreels E., Rieger M., Schaefer M., Meeller-Auer S.,
Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann W., Wedler H., Reinhardt R., Purnelle B.,
Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
A Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
"The genome sequence of Schizosaccharomyces pombe.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 566-855 FROM N.A.
MEDLINE-98382525; PubMed-9714831;
Hoe K.-L., Won M.S., Chung K.-S., Park S.-K., Kim D.-U., Jang Y.-J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YOO O.-J., YOO H.-S.;
"Molecular cloning of gafl, a Schizosaccharomyces pombe GATA factor,
which can function as a transcriptional activator.";
which can function as a transcriptional activator.";
-: FUNCTION: TRANSCRIPTIONAL ACTIVATOR.
-: SUBCELLULAR LOCATION: Nuclear (Probable).
-: SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 855;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                635 CINCQIRITPLWRRSPDGQPLCNACGLFMKINGVVRPLSLKTDVIKKRNR 684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKIDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ch 86.7%; Score 242; DB 1; Length 85
1 Similarity 84.0%; Pred. No. 7.2e-23;
42; Conservative 5; Mismatches 3; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    659 GATA-TYPE.
91776 MW; 3D932F83D2DE6774 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Transcriptional regulatory protein GAT1.
GAT1 OR YFL021W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         510 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AL035076; CAA22647.1; -. EMBL; AL049521; CAB40003.1; -. EMBL; L31001; AAC35593.1; -. HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; 18200679; Znf_GATA.
Pfam; PF00320; GATA. 1.
PRINTS; PR00619; GATAZNFINGER.
SMART; SM00401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        635 6
855 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSFAC; T02831;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nuclear protein.
```

ö

us-09-725-010-2.rsp

9

Page

```
01-FEB-1995
16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coiled coil.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLN3_YEAST
P18494;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ZN FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLN3_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
  SOUR THE TENT OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation—the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                             Coffman J.A., Rai R., Cunningham T., Svetlov V., Cooper T.G.;
"Gatlp, a GATA family protein whose production is sensitive to nitrogen catabolite repression, participates in transcriptional activation of nitrogen-catabolic genes in Saccharomyces cerevisiae.";
Mol. Cell. Biol. 16:847-858(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; DNA-binding; Zinc-finger; Nuclear protein.
ZN_FING 310 334 GATA-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
GATA type zinc finger protein asd4 (Ascus development protein 4).
                                                                                                                                                                                                                                               STRAIN=S288C / AB972;
MEDLINE=95400292; PubMed=7670463;
MUTAKAIN I'., Naltou M., Hagiwara H., Shibata T., Ozawa M.,
Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
Yamazaki M., Tashiro H., Eki T.;
Analysis of the nucleotide sequence of chromosome VI from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariales; Sordariaceae; Neurospora.
NCBL_TaxID=5141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 510;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62D805E42695F35F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85.3%; Score 238; DB 1;
82.0%; Pred. No. 1.4e-22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             426 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND CHARACTERIZATION. MEDLINE-20455773; PubMed=10998244;
                                                                                MEDLINE=96182087; PubMed=8622686;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRINTS; PROO619; GATAZNFINGER. SMART; SM00401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SGD; S0001873; GAT1.
InterPro; IPR000679; Znf_GATA.
Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  510 AA; 56327 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U27344; AAB03516.1; -. EMBL; D50617; BAA09217.1; -. HSSP; P17429; 4GAT. TRANSFAC; T02817; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Neurospora crassa.
                                                                                                                                                                                                              [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                         SEQUENCE FROM N.A.
NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ASD4_NEUCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9HEV5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ASD-4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                               PARCHEMIASLY STANDSTANDS TANDED THAT FUNCTIONS IN SEXUAL - FUNCTION: TRANSCRIPTIONAL RESULATOR THAT FUNCTION: TRANSCRIPTION OF ASD-4 GENE RESULTS IN AGENESIS OF ASCUS AND ASCOSPORE WITH MACROSCOPICALLY NORMAL FRUITING BODY FORMATION. THE GATA-TYPE ZINC FINGER DOMAIN BINDS TO DNA SEQUENCES FROM ITS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minehart P.L., Magasanik B.;
"Sequence and expression of GLN3, a positive nitrogen regulatory gene
of Saccharomyces cerevisiae encoding a protein with a putative zinc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G., Hunicke-Smith S., Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRINTS; PRO0619; GATAZNFINGER.
SMART; SMO0401; ZnF_GATA; 1.
PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
PROSITE; PS50114; GATA_ZN_FINGER_2; 1.
Transcription regulation; DNA-binding; Zinc-finger; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
Feng B., Haas H., Marzluf G.A.; "ASD4, a new GATA factor of Neurospora crassa, displays sequence-specific DNA binding and functions in ascus and ascospore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLN3 ÓR YERÔ40W.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16 CONCATSTTPLWRRDEMGQVLCNACGLFLKLHGRPRPISLKTDVIKSRNR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COILED COIL (POTENTIAL).
4B7DE193050E280E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                               -:- SUBUNIT: HOMOTETRAMER.
-:- SUBCELLULAR LOCATION: Nuclear.
-:- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.1e-19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Rel. 31, Last sequence update) (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 210;
Pred. No. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GATA-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 finger DNA-binding domain.";
Mol. Cell. Biol. 11:6216-6228(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE=92049353; Pubmed=1682800;
                                                                                                                               Biochemistry 39:11065-11073(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nitrogen regulatory protein GLN3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AF319953; AAG45180.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1990 (Rel. 16, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              46532 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75.3%;
76.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                        OWN PROMOTER REGION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GATA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  426 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4GAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF00320;
                                                                                                  development.";
```

ö

us-09-725-010-2.rsp

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PNR OR GATA-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PNR_DROME
P52168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
PNR_DROME
                   REPRESENTATION OF THE PROPERTY OF THE PROPERTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó
                                                                                                                                                                                                                                                     UPSTREAM ACTIVATION SEQUENCE OF GLN1, THE GENE ENCODING GLOTAMINE SYNTHETASE. URE2 MAY CATALYTICALLY INACTIVATE GLN3 IN RESPONSE TO AN INCREASE IN THITRACELLULAR CONCENTRATION OF GLUTAMINE.

-!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and CSD3 genes, a TCP-1-related gene, an open reading frame similar to the DAL80 gene, and a tRNA(Arg)."; Yeast 11:873-883(1995).
                                                                                                                                                                                                                                 NITROGEN
Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C., Petel R.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V., Taylor P., Welt Y., Yelton M., Botstell D., Davis R.W.; Submitted (DEC-1994) to the EMBL/GenBank/DDBU databases.
-!- FUNCTION: POSITIVE NITROGEN REGULATORY PROTEIN. REQUIRED FOR THE ACTIVATION OF TRANSCRIPTION OF A NUMBER OF GENES (INCLUDING THE ALLANTOIN PATHWAY GENES) IN RESPONSE TO THE REPLACEMENT OF GUITAMINE BY GLUTAMATE AS SOURCE OF NITROGEN. BINDS THE NITROGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBL_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA-binding; Zinc-finger;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      72.4%; Score 202; DB 1; Length 730; 75.0%; Pred. No. 6.8e-18; ive 4; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 CINCGIKTIPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKR 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARG/LYS-RICH (BASIC).
P -> G (IN REF. 1).
3159E1844469942E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-NOV-1995 (Rel. 32, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   551 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS00344; GAPA_ZN. FINGER_1; 1.
PROSITE; PS50114; GAPA_ZN_FINGER_2; 1.
Transcription regulation; Activator; DN Nuclear protein; Nitrate assimilation.
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN=S288c / FY1679;
MEDLINE=96090136; PubMed=7483851;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRINTS; PRO0619; GATAZNFINGER. SMART; SM00401; ZnF_GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1995 (Rel. 32, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SGD; S0000842; GLN3.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      730 AA; 79382 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M35267; AAA34645.1; -.
EMBL; U18796; AAB64575.1; -.
PIR; S22280; S22280.
HSSP; P17429; 4GAT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein.
OR YJL110C OR J0806.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF00320; GATA; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSFAC; T02818; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sest Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GZF3_YEAST
P42944;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
GZF3_XFAST
TO ZFAST
TO ZFAST
DT 01-NOV
DT 01-NOV
DT 01-NOV
DT 01-NOV
DT 06-OCT
DE GZF3
DE GZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Burpoen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
15-UDN-2002 (Rel. 41, Last annotation update)
GATA-binding factor-A (Transcription factor GATA-A) (GGATA-A) (Pannier
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIGN: Fructor, con., L. Prents; PRONOS, CARAZNEINGER., PRINTS; PROOGL9; GATAZNEINGER., PROSITE; PSO0344; GATA_ZN_FINGER_1; 1. PROSITE; PSO0114; GATA_ZN_FINGER_2; 1. Transcription regulation; DNA-binding; Zinc-finger; Nuclear protein. Nu FING 131 155 GATA_TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-94139551; PubMed=7916677;
Winick J., Abel T., Leonard M.W., Michelson A.M.,
Chardon-Loriaux I., Holmgren R.A., Maniatis T., Engel J.D.;
"A GATA family transcription factor is expressed along the embryonic dorsoventral axis in Drosophila melanogaster.";
Development 119:1055-1065(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
STRAIN=S288C / FY1679;
MEDIARE-97103755; PubMed=8948101;
MEDIARE-97103755; PubMed=8948101;
"Sequencing analysis of a 40.2 kb fragment of yeast chromosome X reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2, SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes, three remnant delta elements and a Ty4 transposon.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=94139568; PubMed=7916679;
Ramain P., Heitzler P., Haenlin M., Simpson P.;
"Pannier, a negative regulator of achaete and scute in Drosophila,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             72.0%; Score 201; DB 1; Length 551; 74.0%; Pred. No. 7e-18; Live 2; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                        - SUBCELLULAR LOCATION: Nuclear (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SGD; S0003646; GZF3.
InterPro; IPR000679; Znf_GATA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X85021; CAA59384.1; -.
EMBL; Z49385; CAA89405.1; -.
EMBL; X86353; CAA60126.1; -.
HSSP; P17429; 4GAT.
TRANSFĄC; T02827; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best_Local Similarity 74.09
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                 Yeast 12:1471-1474(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00320; GATA; 1
```

us-09-725-010-2.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                          -1- SUBCELLULAR LOCATION: Nuclear.
-1- DEVELOPMENTAL STAGE: IT IS FIRST SEEN IN THE DORSAL PORTION OF THE EMBRYO JUST AFTER CELLULARIZATION AND IS EXPRESSED AT HIGH LEVELS DURING EARLY EMBRYOGENESIS AND AS DEVELOPMENT PROGRESSES HIGH LEVELS ARE SEEN IN THE DORSAL EPIDERMIS.
-1- SIMILARITY: CONTAINS 2 GATA-TYPE ZINC FINGERS.
               transcription factor GATA-1.";
Development 119:1277-1291(1993).
-!- FUNCTION: MAY BE INVOLVED IN DETERMINING DORSAL CELL FATE. IT ACTS
-!- REPRESSOR OF ACHAETE-SCUTE COMPLEX (AS-C) AND IS REQUIRED FOR ITS SPATIAL REGULATION DURING DEVELOPMENT OF THE ADULT PERIPHERAL NERVOUS SYSTEM, AND HENCE FOR THE POSITIONING OF NEURAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA-BINDING.
DNA-BINDING.
DNA-BINDING.
DNA-BINDING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRINTS: PRODGIS, GATAZNEINGER.
SMART; SMO0401; ZnF_GATA.2.
PROSITE; PROD344; GATA_ZN_FINGER_1; 2.
PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
Transcription regulation; Activator; DNA-binding; Zinc-finger;
encodes a zinc finger protein with homology to the vertebrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       226 CTNCGTRITILMERNNDGEPYCNACGLYYKLHGVNRPLAMRKDGIQTRKR 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 198; DB 1; Length 540; Pred. No. 1.7e-17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 CTNCGTKTTPLWRRNPQGQPLCNACGLFLKLHGVVRPLSLKTDVIKKRQR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          E->K: IN PNRD4; ABOLISHES
C->Y: IN PNRD1; ABOLISHES
G->E: IN PNRD2; ABOLISHES
C->S: IN PNRD3; ABOLISHES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -> S (IN REF. 2).
46557C0754BEEB59 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GATA-TYPE 1.
GATA-TYPE 2.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLY-ALA.
POLY-ALA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-GLN.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C->S:
Y -> S
                                                                                                                                                                                                                                                                                                                                                                  EMBL, 568798, AAB29874.1; -.
EMBL, 568999, AAB29874.1; JOINED.
EMBL, 568793, AAB29874.1; JOINED.
EMBL, 568803, AAB29876.2; -.
EMBL, 568803, AAB29876.2; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HSSP; P17679; IGNF.
FlyBase; FBgn0003117; pnr.
InterPro; IPR000679; Znf_GATA.
InterPro; IPR001164; hRIP_like.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         57035 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        71.0%;
66.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00320; GATA; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         540 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nuclear protein.
ZN_FING 169
ZN_FING 226
                                                                                                            PRECURSORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MUTAGEN
MUTAGEN
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MUTAGEN
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

Search completed: January 3, 2003, 19:48:37 Job time : 15 secs